



Dale, James Edward (2014) *Tailoring therapy to individual patient's needs: intensive management of early rheumatoid arthritis using either clinical, laboratory or musculoskeletal ultrasound assessment of disease activity*. PhD thesis.

<http://theses.gla.ac.uk/4991/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Glasgow Theses Service  
<http://theses.gla.ac.uk/>  
theses@gla.ac.uk

**Tailoring Therapy to Individual Patient's Needs**

**Intensive Management of Early Rheumatoid Arthritis**  
**Using Either Clinical, Laboratory or Musculoskeletal Ultrasound**  
**Assessment of Disease Activity**

**By**

**Dr James Edward Dale**  
**MBCChB (Honours), MRCP (UK)**

**Submitted in fulfilment of the requirements for the**  
**Degree of Doctor of Philosophy**  
**University of Glasgow**

**Institute of Infection, Immunity and Inflammation**  
**College of Medical, Veterinary and Life Sciences**  
**University of Glasgow**  
**February 2014**

“Therefore, since I myself have carefully investigated everything from the beginning, it seemed good also to me to write an orderly account”

*Luke Ch1 v 3*

## Declaration

The work presented within this thesis was conducted during my period as a Clinical Research Fellow and Clinical Lecturer within the Institute of Infection, Inflammation and Immunity, College of Medical, Veterinary and Life Sciences, University of Glasgow. The work was funded by a Clinical Academic Fellowship from the Chief Scientist's Office and two project grants from Pfizer UK.

The Targeting Synovitis in Early Rheumatoid Arthritis Study was designed by myself, with guidance from my supervisors Professor Iain McInnes and Dr Duncan Porter. I performed all of the statistical analyses described in Chapters 4-6 and 8. I performed all of the DAS28 assessments, MSUS assessments and was the sole rheumatologist responsible for the care of each participant over the follow-up period. For the comparison between DAS28 and MSUS findings (Chapter 5) Mr David Purves of the Robertson Institute for Biostatistics provided the DAS28 and MSUS data from the TaSER study eCRF whilst I performed the statistical comparisons

The laboratory analysis of the PAXgene RNA samples was conducted by Dr Martin McBride (lecturer), Dr Wai Kwong Lee (research manager), Mrs Iona Donnelly (research associate) and Miss Wendy Crawford (technician) of the Institute of Cardiovascular and Medical Sciences and Mrs Lynn Crawford (senior technician) of the Institute of Infection, Immunity and Inflammation. Dr John McClure (Lecturer) of the Institute of Cardiovascular and Medical Sciences performed the complex statistical analyses described in Chapter 7, whereas I defined all of the grouping variables

The laboratory analysis of serum samples for the Multi-Biomarker Disease Activity score was conducted by laboratory staff from Crescendo Biosciences (South San Francisco). Crescendo Biosciences provided me with the raw data relating to the individual components of the MBDA test and the final MBDA score.

The text contained within this thesis is entirely my own work and has not been submitted in any other form to any other University

James Dale  
September 2013

## Acknowledgements

This thesis has been submitted in my name but would not have been possible without support from a large number of people. I would like to take this opportunity to extend my personal, and heart felt, thanks to the following individuals:

All of the 111 participants, for having the courage and confidence to take part in the TaSER study

My supervisors, Professor Iain McInnes and Dr Duncan Porter, for providing wise counsel, guidance and (occasionally) a little political muscle, throughout, for having sufficient confidence that I would be able to complete all of the necessary tasks and providing sufficient space to pursue my own interests

The invaluable and irreplaceable Sister Anne Stirling for being a wonderful colleague and companion throughout. Without Sister Stirling's experience, commitment and good humour it would not have been possible to complete this study

Rheumatology colleagues throughout NHS Greater Glasgow and Clyde for allowing me to look after their patients, if only for a little while. In particular, Drs John Hunter, Margaret-Mary Gordon, Monica Gupta, Sandeep Bawa, Sarah Saunders, David McCarey, Rajan Madhok, Hilary Wilson, Anne McEntegart, David Marshall, Lisa Hutton, Gillian Roberts, Martin Perry, Fiona McGarry, Nicola Alcorn, Laura McGregor, Islay Morrison, Lucy McGeoch, Catriona Grigor, Meng May Chee, Derek Baxter, Jagtar Nijjar, Jamie Robertson and Sisters Joan Roberts, Janice France, Moira McDonald, Anne Simmons, Liz McIvor and Victoria Toner

Drs Martin McBride and John McClure for their technical assistance in conducting the gene expression analysis and their patience with me coming to terms with the new techniques

Mr Alan Stevenson, Mr David Purves and Dr Sharon Kean of the Robertson Institute of Biostatistics for devising and curating the TaSER eCRF and assisting with my frequent requests for additional data

Dr Nadine Defranoux and her colleagues at Crescendo Biosciences in San Francisco for agreeing to perform the MBDA analysis

Drs Debbie Turner and Anna Ciechomska for introducing me to musculoskeletal ultrasound and not minding when my training requirements caused their clinics to run late

Professor Phillip Conaghan, University of Leeds, for advice on the components of the 14 joint MSUS set and MSUS definitions of active disease

Mrs Caron Paterson and her team for cataloguing the TaSER blood samples and arranging their safe transport to California

The staff of the MRI department at Glasgow Royal Infirmary and the NHS GGC Central Booking Office

Most of all, I must extend my deepest and warmest thanks and love to my wonderful wife Arlene and our two sons Zack and Archie, for allowing me to spend so much of our family time conducting the study and writing this thesis. The value of their patience, humour and love cannot be described

# Table of Contents

<b>Declaration</b>	<b>3</b>
<b>Acknowledgements</b>	<b>4</b>
<b>Table of Contents</b>	<b>6</b>
<b>Index of Tables</b>	<b>11</b>
<b>Index of Graphs</b>	<b>14</b>
<b>Publications</b>	<b>17</b>
<b>Abstracts</b>	<b>17</b>
<b>Summary</b>	<b>18</b>
<b>1. Introduction</b>	<b>21</b>
<b>1.1 Clinical Features</b>	<b>22</b>
1.1.1 Diagnosis	22
1.1.2 Clinical Course	25
1.1.3 The Pre-symptomatic Stages of Rheumatoid Arthritis	35
<b>1.2 Current Approaches to the Management of Early Rheumatoid Arthritis</b>	<b>37</b>
1.2.1 Core Principles	37
1.2.2 DMARD Treatment Regimens	42
<b>1.3 Global Disease Activity Assessment</b>	<b>48</b>
1.3.1 Clinical examination	48
1.3.2 Acute Phase Reactants	50
1.3.3 Composite Disease Activity Measures	51
<b>1.4 The Role of Musculoskeletal Ultrasound in the Assessment of Rheumatoid Arthritis</b>	<b>56</b>
1.4.1 Persistent Disease Activity Despite Clinical Remission	56
1.4.2 Inappropriate Elevation of Disease Activity Measure	59
<b>1.5 The Significance of Inflammatory Lesions Identified by Musculoskeletal Ultrasound</b>	<b>62</b>
1.5.1 MSUS appearances of synovitis	63
1.5.2 Power Doppler signal corresponds to active synovitis	66
1.5.3 Implication of MSUS Disease Activity Assessment	70
<b>1.6 Assessment of Rheumatoid Arthritis Disease Activity and Prognosis Using Multilevel Biomarker Profiles</b>	<b>71</b>
1.6.1 Expression Analysis	72
1.6.2 Metabolomics	78
<b>1.7 Objectives</b>	<b>81</b>
<b>1.8 Hypotheses</b>	<b>82</b>
<b>2. Methods</b>	<b>83</b>
<b>2.1 Introduction</b>	<b>84</b>
<b>2.2 Identification of Study Cohort</b>	<b>84</b>
2.2.1 Screening Arrangements	84
2.2.2 Inclusion Criteria	86
2.2.3 Exclusion Criteria	88

2.2.4	Sample Size Estimation	90
<b>2.3</b>	<b>DMARD Treatment Protocol</b>	<b>91</b>
2.3.1	Overview of DMARD Escalation Protocol	91
2.3.2	<i>Clinical Factors Relating to Individual Agents</i>	95
2.3.3	<i>The Evidence Supporting the DMARD Escalation Regimen</i>	99
<b>2.4</b>	<b>DMARD Escalation Thresholds</b>	<b>107</b>
2.4.1	Randomisation Process	107
2.4.2	DMARD Escalation Thresholds for Each Group	108
<b>2.5</b>	<b>Outcome Measures</b>	<b>121</b>
2.5.1	Clinical Outcome Measures	121
2.5.2	Laboratory Outcome Measures	123
2.5.3	Radiological Outcome Measures	123
2.5.4	Data Storage and Management	126
<b>2.6</b>	<b>Statistical Analysis</b>	<b>127</b>
2.6.1	Description of Baseline Characteristics	127
2.6.2	Assessment of Impact of MSUS Upon DMARD Escalation Decision Making	128
2.6.3	Description of Treatment Intensity	130
2.6.4	Description of Treatment Response	131
2.6.5	Description of Change in Radiological Outcomes	134
2.6.6	Adverse Event Rates	135
<b>2.7</b>	<b>Biomarker Analysis</b>	<b>138</b>
2.7.1	Principles of Biomarker Analysis	138
2.7.2	Transcriptomic Analysis	141
<b>3.</b>	<b><u>Practical Considerations, Unexpected Developments and Research Design Amendments</u></b>	<b>153</b>
<b>3.1</b>	<b>Dates</b>	<b>154</b>
<b>3.2</b>	<b>Research Funding</b>	<b>155</b>
<b>3.3</b>	<b>NHS Research Ethics and Management Approvals</b>	<b>155</b>
<b>3.4</b>	<b>Participant Screening, Consent and Recruitment Rates</b>	<b>156</b>
<b>3.5</b>	<b>Musculoskeletal Ultrasound Training</b>	<b>158</b>
<b>3.6</b>	<b>Identification of Collaborator for Transcriptomic Analysis</b>	<b>159</b>
<b>3.7</b>	<b>Recruitment to Scottish Early Rheumatoid Arthritis Inception Cohort</b>	<b>160</b>
<b>3.8</b>	<b>Amended Analysis Plan</b>	<b>161</b>
3.8.1	Interim MSUS Data Analysis Plan	161
3.8.2	Amended Transcriptomic Expression Analysis Plan	163
<b>4.</b>	<b><u>The TaSER Study Cohort</u></b>	<b>164</b>
<b>4.1</b>	<b>Identification of Participants</b>	<b>165</b>
4.1.1	Referral Sources	165
4.1.2	Screening Process	165
<b>4.2</b>	<b>Baseline Characteristics of Study Cohort</b>	<b>167</b>
4.2.1	Fulfilment of RA classification criteria	169
4.2.2	Distribution of Joint Involvement at Presentation	172
<b>4.3</b>	<b>Discussion</b>	<b>172</b>
4.3.1	Baseline Clinical Features	172
4.3.2	Baseline Autoantibody Status	174
4.3.3	Fulfilment of Existing RA Classification Criteria	176
4.3.4	Characteristics of other recent early RA treatment strategy studies	178
<b>5.</b>	<b><u>Impact of Musculoskeletal Ultrasound on DMARD Escalation Decisions</u></b>	<b>180</b>
<b>5.1</b>	<b>Fulfilment of Indication Criteria for MSUS</b>	<b>181</b>
	Fulfilment of MSUS Indication Criteria by Assessment Group	181

<b>5.2</b>	<b>Clinical Disease Activity Assessment Findings in MSUS Assessment Group</b>	<b>183</b>
5.2.1	Mean Values of DAS28 Component Variables	183
5.2.2	Frequency of Clinical Disease Activity States in MSUS Assessment Group by Month	184
<b>5.3</b>	<b>Potential workload implications of MSUS Assessment</b>	<b>185</b>
5.3.1	Frequency of each MSUS indication and likelihood of identifying active disease	185
5.3.2	Relationship of frequency of MSUS assessment to incidence of active disease	187
5.3.3	Indication and findings of first MSUS assessment	189
<b>5.4</b>	<b>Change in MSUS PD signal findings over follow-up period</b>	<b>189</b>
5.4.1	Overall change in MSUS findings	189
5.4.2	Time course assessment of changes in PD findings	191
<b>5.5</b>	<b>MSUS PD signal findings by joint area</b>	<b>194</b>
5.5.1	Frequency of positive MSUS assessments using different joint sets	195
<b>5.6</b>	<b>Agreement between DAS28 and MSUS disease activity assessments</b>	<b>196</b>
5.6.1	Definitions (reprise)	196
5.6.2	MSUS disease activity assessment findings during DAS28 LDAS	197
5.6.3	MSUS disease activity assessment findings during DAS28 remission	197
5.6.4	MSUS disease activity assessment findings during moderate disease activity and minimal clinical synovitis	197
5.6.5	Overall agreement between DAS28 and MSUS disease activity assessments	198
<b>5.7</b>	<b>Change in MSUS synovial hypertrophy findings over follow-up period</b>	<b>199</b>
<b>5.8</b>	<b>Discussion</b>	<b>201</b>
5.8.1	Selection of patients for MSUS assessment	201
5.8.2	Proposed MSUS joint set and definitions of active RA	204
5.8.3	Workload implications of regularly performing MSUS disease activity assessment	206
5.8.4	Impact of MSUS assessment upon DMARD escalation decisions	208
<b>6.</b>	<b>Impact of MSUS Disease Activity Assessment upon Clinical Outcomes</b>	<b>213</b>
<b>6.1</b>	<b>Clarification of disposition of study participants</b>	<b>214</b>
<b>6.2</b>	<b>Comment on the statistical analysis</b>	<b>215</b>
<b>6.3</b>	<b>Overall change in ACR core set variables</b>	<b>215</b>
<b>6.4</b>	<b>Change in DAS44</b>	<b>217</b>
6.4.1	Mean Change in DAS44 from baseline	217
6.4.2	DAS44 at each time point	218
6.4.3	EULAR response rates	219
6.4.4	Rate of Low Disease Activity and Remission	220
<b>6.5</b>	<b>Change in HAQ</b>	<b>223</b>
<b>6.6</b>	<b>Change in Other ACR Core Set Variables</b>	<b>224</b>
6.6.1	44 Swollen Joint Count	224
6.6.2	Change in Ritchie Articular Index	225
6.6.3	Change in ESR	225
6.6.4	Change in CRP	226
6.6.5	Change in Global Health 10cm VAS	227
6.6.6	Change in Pain 100mm VAS	227
<b>6.7</b>	<b>Impact of gender on treatment response and remission rates</b>	<b>228</b>
6.7.1	Impact of gender on DAS44 and HAQ – whole cohort	228
6.7.2	Impact of gender on DAS44 and HAQ – by assessment group	229
<b>6.8</b>	<b>Multiple Regression Analyses</b>	<b>232</b>
<b>6.9</b>	<b>Discussion</b>	<b>233</b>
<b>7.</b>	<b>The Relationship Between RA Phenotype and Gene Expression Profile</b>	<b>238</b>

<b>7.1</b>	<b>Clinical Features of Gene Expression Analysis Cohort</b>	<b>239</b>
<b>7.2</b>	<b>Baseline Phenotypic Groupings</b>	<b>241</b>
7.2.1	Gender	241
7.2.2	Rheumatoid Factor Status	244
7.2.3	Anti-CCP Antibody Status	245
7.2.4	Baseline Erosive Status	245
7.2.5	Smoking Status	246
<b>7.3</b>	<b>Gene Expression Profiles as Measures of Global Disease Activity</b>	<b>246</b>
7.3.1	Baseline Gene Expression Profile and Baseline Disease Activity – DAS28 Quartiles	247
7.3.2	Three Month Gene Expression Profile and Three Month Disease Activity – DAS28 Quartiles	247
7.3.3	Three Month Gene Expression Profile and Three Month Disease Activity – DAS28 Thresholds	247
7.3.4	Three Month Gene Expression Profile and Clinical Threshold for DMARD Escalation	248
<b>7.4</b>	<b>Baseline Gene Expression Profile as Predictor of Clinical Course</b>	<b>248</b>
7.4.1	Baseline Gene Expression Profile and Three Month Threshold for DMARD Escalation	248
7.4.2	Baseline Gene Expression Profile and Three Month Disease Activity Level – DAS28 Quartiles	249
7.4.3	Baseline Gene Expression Profile and Clinical Course – Mean DAS28 between 0 and 12 months	249
7.4.4	Dynamic Changes in Gene Expression Profile in Relationship to Changes in Disease Activity	250
<b>7.5</b>	<b>Additional Unplanned Phenotypic Groupings</b>	<b>250</b>
7.5.1	Correcting for Influence of Gender on Gene Expression Profiles	250
7.5.2	Dynamic Change in Gene Expression Profile Between Baseline and 3 Months	251
7.5.3	Incorporation of MSUS Disease Activity Assessment Findings into Phenotypic Groupings	252
7.5.3	Comparison of Differences Between Baseline and 3 Month Gene Expression Profiles in Patients Who Achieved DAS28 Remission After 3 Months	253
7.5.4	Relationship of Interferon Response Gene Expression to Disease Activity Levels	254
<b>7.6</b>	<b>Discussion</b>	<b>262</b>
7.6.1	Relationship between baseline phenotypic characteristics and gene expression profile	263
7.6.2	Relationship between RA disease activity state and gene expression profile	268
7.6.3	Relationship between treatment response and baseline gene expression profile	273
<b>8.</b>	<b>Disease Activity Assessment Using a Multi-Biomarker Disease Activity Test</b>	<b>278</b>
<b>8.1</b>	<b>Introduction</b>	<b>279</b>
8.1.1	The Multi-Biomarker Disease Activity Test	279
8.1.2	Objective	282
<b>8.2</b>	<b>Methods</b>	<b>282</b>
<b>8.3</b>	<b>Results</b>	<b>283</b>
8.3.1	Change in DAS28-ESR and MBDA Score	284
8.3.2	Relationship of DAS28-ESR to MBDA Score	285
8.3.3	Percentage agreement between DAS28-ESR and MBDA	287
8.3.4	Percentage agreement between DAS28-ESR and MSUS assessment	289
8.3.5	Relationship between MBDA score and subsequent disease activity	291
8.3.6	Change in MBDA score by randomisation group	295
8.3.7	Change in MBDA score in response to anti-TNF $\alpha$ blocking therapy	296
<b>8.4</b>	<b>Discussion</b>	<b>297</b>
<b>9.</b>	<b>Final Discussion and Considerations</b>	<b>303</b>
	<b>Considerations</b>	<b>309</b>

<b>9.1 The TaSER study</b>	<b>309</b>
Length of follow-up period	309
Group Size	309
Baseline characteristics of assessment groups	310
Proposed MSUS definition of active disease	310
Choice and performance of clinical outcome measures	311
Radiographic outcomes	312
Adverse events rate	313
DMARD escalation protocol	313
Bias of investigator	315
Incomplete DAS28 vs MSUS dataset	315
Absence of MSUS data for DAS28 group	316
<b>9.2 The Gene Expression Analysis</b>	<b>317</b>
Choice of target tissue	317
Potential impact of laboratory technique on gene expression	318
Choice of microarray platform	320
Choice of statistical analysis technique	320
Choice of comparator group descriptors	321
Validation	323
<b>9.3 The Multi-Biomarker Disease Activity Test Analysis</b>	<b>323</b>
Temporal relationship of disease activity to DAS28-ESR and MBDA test	323
Use of DAS28-ESR rather than DAS28-CRP	324
Sensitivity of MBDA in relation to alternative disease activity measures	324
MSUS joint set	325
<b>10. Conclusions</b>	<b>326</b>
<b>Areas for Future Work</b>	<b>331</b>
<b>References</b>	<b>334</b>
<b>Appendix A - Abbreviations</b>	<b>365</b>
<b>Appendix B – Recruitment Poster</b>	<b>366</b>
<b>Appendix C – Patient Information Leaflet</b>	<b>367</b>
<b>Appendix D – Patient Consent Form</b>	<b>369</b>
<b>Appendix E – Recording Instrument for MSUS Findings</b>	<b>370</b>
<b>Appendix F: The Health Assessment Questionnaire</b>	<b>371</b>
<b>Appendix G: The EQ-5D Questionnaire</b>	<b>374</b>
<b>Appendix H: Materials for Transcriptomic Analysis</b>	<b>376</b>

## Index of Tables

Table 1: Summary of components of RA composite disease activity measures and scoring ranges	53
Table 2: Summary of sample size calculation parameters for TICORA and TEAR studies	91
Table 3: Summary of DMARD escalation steps in early RA treatment strategy studies	101
Table 4: Timing of collection of each outcome measure	121
Table 5: Derivation of EULAR response using DAS44 thresholds	122
Table 6: Potential range of scores for RAMRIS components for wrist and 2 <sup>nd</sup> -5 <sup>th</sup> MCPj separately and combined	125
Table 7: Baseline characteristics and proposed statistical analysis methods	128
Table 8: Potential range of scores for plain xray erosions and joint space narrowing in the hands and feet when graded using the Sharp / Van der Heijde Score	135
Table 9: Vacutainer set and draw order	138
Table 10: Illustration of biomarker sample labelling system for sample (T00X – $\alpha$ – $\beta$ )	138
Table 11: Initial blood sample handling procedures	139
Table 12: Summary of phenotypic grouping structures and corresponding mRNA expression profile sampling time points	151
Table 13: Original source of referral for research participants	165
Table 14: Follow-up arrangements for research participants	165
Table 15: Frequency of reasons why patients with RA (n=59) did not participate in research	166
Table 16: Frequency of alternative diagnoses in patients (n = 113) attending for screening review	166
Table 17: Baseline demographic and disease related characteristics of whole research cohort, DAS28 and MSUS assessment groups	168
Table 18: Frequency of anti-CCP and/or rheumatoid factor positivity for the whole research cohort	169
Table 19: Fulfilment of 1987 ACR Classification Criteria – during baseline assessment	170
Table 20: Fulfillment of 1987 ACR Classification Criteria – total score at baseline assessment	170
Table 21: Fulfilment of 2010 ACR-EULAR Classification Criteria for RA – during baseline assessment.	171
Table 22: Fulfillment of 2010 ACR-EULAR Classification Criteria for RA – total score at baseline assessment.	171
Table 23: Frequency and distribution of clinically evident synovitis at presentation	172
Table 24: Comparison of rates of rheumatoid factor and anti-CCP antibody positivity between TaSER study cohort and NHSGGC early arthritis cohort	175
Table 25: Comparison of rates of rheumatoid factor and anti-CCP antibody positivity between TaSER study cohort and Swedish early arthritis cohort	176

Table 26: Baseline RA disease characteristics and disease activity levels for patients recruited to the TASER, TICORA, TEAR (UK) and SWEFOT studies of RA treatment strategies	179
Table 27: Disease activity states and relationship to DMARD escalation decisions	181
Table 28: Frequency (percentage) that DAS28 assessments fulfilled different disease activity states between 3 and 18 months.	182
Table 29: Mean (SD) values of DAS28 components for MSUS assessment group participants between 3 and 18 months follow-up	184
Table 30: Number of first MSUS assessments performed for each MSUS indication and number (percentage) that identified active disease	189
Table 31: Relative risk of subsequent MSUS assessments identifying active disease following consecutive negative MSUS assessments	194
Table 32: Frequency (percentage) of positive MSUS assessments using different joint sets	196
Table 33: Agreement (green boxes) and disagreement (red boxes) between DAS28 and MSUS assessments of global disease activity.	198
Table 34: Agreement (green boxes) and disagreement (red boxes) between DAS28 and MSUS assessments of global disease activity.	198
Table 35: Agreement (green boxes) and disagreement (red boxes) between DAS28 and MSUS assessments of global disease activity.	198
Table 36: Rate of positive MSUS assessments in relation to number of MSUS assessments performed per patient	207
Table 37: Range of skew across the whole follow-up period for each ACR core set variable	215
Table 38: Overall change in each ACR Core Set Variable between baseline and 18 months	216
Table 39: Odds of participants within DAS28 group attaining DAS44 remission on separate occasions compared to MSUS group	222
Table 40: Odds of participants within DAS28 group attaining DAS44 remission on consecutive occasions compared to MSUS group	223
Table 41: Baseline features of transcriptomic analysis cohort in comparison to the whole research cohort.	240
Table 42: Disease activity measures at baseline, 3 months and 18 months for the transcriptomic analysis cohort.	240
Table 43: DAS44, and change from baseline of DAS44, at baseline, 3 months and 18 months for transcriptomics analysis cohort and whole research cohort.	241
Table 44: Number (percentage) of transcriptomic analysis cohort attaining DAS44-defined disease activity states at each time point	241
Table 45: Genes exhibiting differential expression in baseline samples between female and male RA patients. Ranked by adjusted p value	243
Table 46: Genes exhibiting differential expression between baseline samples of rheumatoid factor positive and negative patients	245
Table 47: Genes exhibiting differential expression in baseline samples between current, former and never smokers	246

Table 48: Characteristics of a single gene which was down regulated in participants who failed to achieve 3 month DAS28 < 3.2	248
Table 49: Genes which were up-regulated in baseline samples of patients with 3 month DAS28 greater than 3.2	249
Table 50: Genes which were differentially expressed in baseline samples of female participants grouped by rheumatoid factor status	251
Table 51: Genes which underwent changes of expression between baseline and 3 months	253
Table 52: Disease activity category definitions based upon DAS-CRP and MBDA	280
Table 53: Comparison of thresholds for MBDA, DAS28 and MSUS definitions of RA disease activity	283
Table 54: Comparable dichotomous disease activity states as defined by DAS28-ESR and MBDA	289
Table 55: Summary of MSUS disease activity findings by MBDA disease activity group.	290
Table 56 and Graph 80 – MBDA disease activity definition thresholds for this cohort – estimated by interpolating linear regression of correlation between all DAS28-ESR and MBDA	299

## Index of Graphs

Graph 1: Frequency of MSUS indication by month – MSUS assessment group only	183
Graph 2: Percentage occurrence of each disease activity state by month – MSUS assessment group	185
Graph 3 – DAS28<2.6 - Frequency of MSUS assessments and frequency of positive MSUS assessments	186
Graph 4 - DAS28 < 3.2 - Frequency of MSUS assessment and frequency of positive MSUS assessments	186
Graph 5 - $3.2 \leq \text{DAS28} < 5.1$ and $\text{SJC} < 2$ – Frequency of MSUS assessment and frequency of positive MSUS assessments	186
Graph 6 – Relationship between frequency of MSUS assessment and frequency of positive MSUS assessments	187
Graph 7 – Percentage of MSUS assessments that identified active disease by MSUS indication and month	187
Graph 8 – Distribution of number of MSUS assessments conducted per patient over follow-up period	188
Graph 9 – Scattergraph demonstrating the relationship between the number of MSUS assessments performed per patient and the number of assessments that identified active disease	188
Graph 10 – Box and whisker plot of dispersal of PD joint counts recorded each month – all results.	190
Graph 11 – Box and whisker plot of dispersal of total PD score recorded each month	190
Graph 12 – Number of positive MSUS assessments observed for each month after first MSUS assessment	191
Graph 13 - Total PD joint count score for each month after first MSUS assessment (month 0)	192
Graph 14 – Risk of subsequent MSUS assessment identifying active disease following consecutive negative assessments.	194
Graph 15 – Frequency and distribution of positive PD findings by joint area	195
Graph 16 - Percentage agreement between DAS28 and MSUS disease activity assessments based upon different MSUS PD joint count thresholds	199
Graph 17 – Mean total synovial hypertrophy score (black line) and joint count (dotted line) by month	200
Graph 18 – Box and whisker plot of range of values for total synovial hypertrophy score for each month	201
Graph 19 – Mean change from baseline of DAS44	217
Graph 20 – Median DAS44 score at each assessment time point	218
Graph 21 – DAS28 Group – percentage of participants fulfilling EULAR response criteria	219
Graph 22 - MSUS Group – percentage of participants fulfilling EULAR response criteria	219

Graph 23 - Percentage of participants within each group achieving DAS44 LDAS (DAS44 < 2.4) between follow-up months 3 to 18	220
Graph 24 – Percentage of participants within each group achieving DAS44 remission (DAS44 <2.4) between follow-up months 3 to 18	221
Graph 25 – Percentage of participants attaining ACR/EULAR remission follow-up period	222
Graph 26 – Median HAQ at each assessment time point	223
Graph 27 - Mean HAQ at each assessment time point	224
Graph 28 – Median 44SJC at each assessment time point	225
Graph 29 – Median Ritchie Articular Index at each assessment time point	225
Graph 30 – Median ESR at each assessment time point	226
Graph 31 – Median CRP at each assessment point	226
Graph 32 – Median Global Health 10cm Visual Analogue Score at each assessment time point	227
Graph 33 – Median Pain VAS at each assessment time point	227
Graph 34 – Median DAS44 over follow-up period – segregated by gender	228
Graph 35 - Percentage of participants attaining DAS44 remission, stratified by gender	229
Graph 36 – Median HAQ at each assessment time point, segregated by gender	229
Group 37 – DAS28 group: median DAS44 at each assessment time point, segregated by gender	230
Graph 38 – MSUS group – median DAS44 at each assessment time point, segregated by gender	230
Graph 39 – Median DAS44 for female participants, stratified by assessment group	231
Graph 40 – Median DAS44 for male participants, stratified by assessment group	231
Graph 41 – MSUS group: Median HAQ score at each assessment time point, segregated by gender	232
Graph 42 – DAS28 group: Median HAQ score at each assessment time point, segregated by gender	232
Graph 43 – IRG T3:T0 ratio according to 3 month clinical response.	256
Graph 44 – IRG T3:T0 ratio between patients with low disease activity (DAS28 <3.2) and moderate disease activity or higher (DAS28 >3.2) at 3 months	257
Graph 45 – Association between 3 month IRG T3:T0 ratio and clinical response at 6 months.	258
Graph 46 – IRG T3:T0 ratio between patients with low disease activity (DAS28 <3.2) or moderate disease activity or higher (DAS28 >3.2) at 6 months	258
Graph 47 – Association between 3 month IRG T3:T0 ratio and clinical response at 12 months.	259
Graph 48 - IRG T3:T0 ratio between patients with low disease activity (DAS28 <3.2) or moderate disease activity or higher (DAS28 >3.2) at 12 months	259
Graph 49 – Baseline IRG score in relation to 3 month clinical response	261

Graph 50 – Baseline IRG score in relation to 3 month clinical response	261
Graph 51 – Baseline IRG score in relation to 12 month clinical response	262
Graph 52 – Baseline IRG score in relation to 12 month clinical response	262
Graph 53 – Mean DAS28 and MBDA values at baseline, 3 months and 18 months	284
Graphs 54 – Distribution of disease activity state classification by DAS28-ESR (54A) and MBDA (54B) at baseline, 3 months and 18 months	285
Graph 55 – Relationship between DAS28-ESR and MBDA – all sets of paired DAS28-ESR and MBDA values pooled from each time point	286
Graph 56 – Relationship between change in DAS28 and corresponding change in MBDA.	287
Graph 57 – Percentage exact agreement between DAS28-ESR and MBDA at each time point and pooled for all available results	287
Graph 58 – Distribution of MBDA disease activity findings when participants were grouped by DAS28-ESR disease activity state. (Results pooled for all time points)	288
Graph 59 – Percentage agreement of dichotomous disease activity assessments by DAS28-ESR and MBDA at each time point.	289
Graph 60 – Musculoskeletal ultrasound findings when participants are stratified by MBDA disease activity definition.	290
Graph 61 – Scattergraph depicting relationship between baseline MBDA and mean DAS28-ESR between baseline and follow-up month 12	292
Graph 62 – Outcome of 3 month MBDA disease activity categorisation based upon baseline MBDA disease activity state	293
Graph 63 – Outcome of 18 month MBDA disease activity categorisation based upon 3 month MBDA disease activity state	294
Graph 64 – Median DAS28-ESR by randomisation group for participants included within the MBDA analysis	295
Graph 65 – Median MBDA score by randomisation group at baseline, 3 months and 18 months.	296
Graph 66 – Median MBDA score in relation to etanercept therapy	296
Graph 67 – Distribution of MBDA-defined disease activity in relation to etanercept therapy	297
Table 56 and Graph 68 – MBDA disease activity definition thresholds for this cohort – estimated by interpolating linear regression of correlation between all DAS28-ESR and MBDA	299
Graphs 69A + 69B – Relationship of baseline MBDA disease activity to 18 month DAS28-ESR and MBDA disease activity grouping	299

## Publications

Porter D, **Dale J**, Sattar N  
How low to aim in RA? Learning from other disciplines  
Ann Rheum Dis 2013; in submission

**Dale J**, Purves D, Macconnachie A, McInnes I, Porter D  
Tightening Up? Impact of musculoskeletal ultrasound disease activity assessment on treatment decisions in rheumatoid arthritis patients treated using a treat to target strategy  
Arth Care Research 2013; accepted pending revision

Crilly A, Burns E, Nickdel M, Lockhart JC, Perry M et al  
PAR2 expression in peripheral blood monocytes of patients with rheumatoid arthritis  
Ann Rheum Dis 2012 (71); p 1049-1054

**Dale J**, Porter D.  
Optimising the strategy of care in early rheumatoid arthritis  
Best Practice and Clinical Research Rheumatology 2010 (24); p 443-455

## Abstracts

**Dale J**, Stirling A, McInnes I, Porter D  
Targeting ultrasound remission in early rheumatoid arthritis – the results of the TaSER Study  
Accepted for oral presentation – American College of Rheumatology Meeting, October 2013

Robertson J, **Dale J**, Sattar N, Porter N  
Aggressive DMARD therapy elevates HDL-cholesterol and lowers the atherogenic index in the TaSER study  
Poster presentation – European League Against Rheumatism Meeting, June 2013

**Dale J**, Purves D, Macconnachie A, McInnes I, Porter D  
Tightening Up? Musculoskeletal Ultrasound Could Further Individualise Treatment Decisions in Early Rheumatoid Arthritis Patients Treated by a Step-up DMARD Escalation Regimen  
Oral presentation – American College of Rheumatology Meeting, November 2012

**Dale J**, Porter D  
Results of the Scottish Early Arthritis Treatment Online Survey  
Oral presentation – Scottish Society of Rheumatology Autumn Meeting, October 2010

**Dale J**, Saunders S, Wilson H, Thomas M, Porter D  
Correlation between anti-cyclic citrillinated protein antibody titres and baseline measures of disease activity in early rheumatoid arthritis  
Oral presentation – Scottish Society of Rheumatology Spring Meeting, June 2010

**Dale J**, Gupta M, Porter D  
The use of musculoskeletal ultrasound to assess disease activity in rheumatoid arthritis: a feasibility exercise  
Poster presentation – British Society of Rheumatology Annual Meeting, April 2009

## Summary

**Background:** Outcomes in the management of early rheumatoid arthritis (RA) have been significantly improved through the use of composite disease activity measures (such as the DAS28) and aggressive DMARD escalation until a lower disease activity target has been achieved. Imaging studies suggest that the DAS28 may be insensitive to low levels of subclinical active disease that is associated with an increased risk of flare and progressive joint damage. Further, in some cases, elevations of the DAS28 may not necessarily be related to ongoing active synovitis. In both instances, relying upon the DAS28 assessment alone may lead to patients being considered for an inappropriate treatment decision since, patients with active subclinical disease may not be considered for further DMARD escalation whilst patients with non-inflammatory causes of DAS28 elevation may be offered additional DMARD therapy that is either ineffective or potentially toxic. There is emerging evidence that musculoskeletal ultrasound (MSUS), gene expression profiles and inflammatory protein microarrays might provide useful additional disease activity information that allows clinicians to reach a treatment decision that is targeted at an individual patient's specific needs

### Objectives:

1. To determine whether using MSUS assessment of global disease activity in addition to the DAS28 produces significantly better short-medium term clinical and radiological outcomes
2. To determine whether grouping early RA patients by either RA phenotype or disease activity level is associated with evidence of differential gene expression between the comparator groups
3. To determine the degree of correlation and agreement between the Multi-Biomarker Disease Activity (MBDA) test, the DAS28 and a MSUS disease activity assessment

### Methods

111 patients with either clinical diagnoses of early RA (symptom duration < 1 year) or anti-CCP antibody positive inflammatory arthritis were recruited to the Targeting Synovitis in Early Rheumatoid Arthritis (TaSER) study. Clinical consultations occurred monthly for 18 months and all participants were treated using the same step-up DMARD-biologic escalation protocol. Participants were randomised to either a DAS28 or MSUS assessment group. In the DAS28 group, DMARD therapy was escalated until DAS28 low disease activity (LDAS – DAS28 <3.2) had been achieved. In the MSUS group, MSUS assessment was indicated for instances of DAS28 LDAS or DAS28 moderate disease activity ( $3.2 \leq \text{DAS28} < 5.1$ ) with minimal clinical synovitis (28SJC  $\leq 1$ ). During MSUS assessment, the bilateral radiocarpal, index and middle MCP, index and middle PIP and 2<sup>nd</sup> and 5<sup>th</sup> MTP joints were examined for the presence of gray scale synovial hypertrophy and Power Doppler (PD) signal. Active disease was defined as the presence of grade 1 or higher PD signal in 2 or more joints. DMARD therapy was not changed if there had been significant escalation within the preceding 3 months. Intra-articular and intra-muscular corticosteroid injections were administered generously during periods of active disease.

Blinded clinical outcomes were collected at baseline and every 3 months until study completion. Plain x-rays of hands and feet and MRI of the dominant wrist and hand were performed at baseline and study completion and will be graded by 2 independent radiologists who are blinded to participant's randomisation group. Primary outcomes comprised: 1. mean change in DAS44 from baseline and 18 months, 2. mean change in MRI RAMRIS erosion score between baseline and 18 months. Secondary outcome measures included: between group comparisons of the DAS44 and ACR-EULAR remission rates, EULAR response criteria, HAQ, EURO-QoL 5D, CRP, ESR, 10cm pain visual analogue score, mean change in plain x-ray Sharp score (van der Heijde modification) and mean change in MRI RAMRIS synovitis and bone marrow oedema scores.

79 Participants donated additional blood samples for nested biomarker analysis at baseline, follow-up months 3 and 18. Baseline and 3 month PAXgene RNA samples were analysed with the assistance of the Systems Biology Group, Institute of Cardiovascular and Medical Sciences, University of Glasgow using an Illumina HumanHT-12v4 Beadchip microarray. Baseline, 3 month and 18 month serum samples were analysed by Crescendo Biosciences using their in house MBDA microarray. Additional whole blood, serum and plasma samples remain available for future polyomic analyses. For the gene expression analysis, participants were segregated into comparator groups based upon baseline and 3 month RA phenotypic and disease activity data. Comparator groups were intended to represent common clinical scenarios. Between group comparisons of gene expression were conducted in the R software package using the Linear Models for Microarray Data (Limma) plug-in. An adjusted p value <0.05 was considered to represent evidence of differential gene expression. For the MBDA analysis, the degree of correlation (Spearman's rank correlation) between DAS28 and MBDA score was calculated at each time point and for all time points pooled together. The percentage agreement between MBDA, DAS28 and MSUS disease activity state categorisations was also calculated.

## **Results**

111 participants were recruited and 101 (91%) completed follow-up. 95 (86%) participants fulfilled 1987 ACR RA classification criteria and 107 (96%) fulfilled 2010 ACR-EULAR RA classification criteria. The presenting features appeared typical of an early RA cohort and, excepting gender, there were no statistical differences in baseline characteristics between the groups

414 MSUS assessments were performed, 369 MSUS assessments coincided with DAS28 LDAS, of which 92 (25%) identified active synovitis. 271 MSUS assessments coincided with DAS28 remission, of which 66 (24%) identified active synovitis. 45 MSUS assessments coincided with DAS28 moderate disease activity of which 15 (33%) identified active synovitis. Overall 71% of paired DAS28 and MSUS assessments agreed on the disease activity state

MSUS-driven DMARD escalation was not associated with significant improvements in clinical outcomes. Both groups experienced a similar mean change in DAS44 between baseline and 18 months (DAS28 -2.51 vs MSUS -2.76, p 0.39). There were no statistically significant between

group differences in the ACR core set variables at any of the time points, nor their mean change from baseline. Over the follow-up period, the MSUS assessment group demonstrated incremental increases in the proportion of participants with EULAR good responses and DAS44 remission and a significantly higher rate of DAS44 remission at study completion (DAS28 44% vs MSUS 65%,  $p=0.045$ ). The impact of MSUS-driven DMARD escalation on radiological outcomes, medium-long term outcomes and adverse event rates remains to be determined.

At baseline, gender (61 genes), RhF status (5 genes) and current smoking (1 gene) were associated with evidence of differential gene expression. The expression patterns of 19 genes changed following commencement of DMARD monotherapy. However, it was not possible to demonstrate evidence of differential gene expression in relation to disease activity level or phenotypic extremes at either time point. Up-regulation of 3 genes at baseline was associated with requiring DMARD escalation at 3 months. Otherwise, baseline gene expression was not predictive of 3 month disease activity state nor disease course over 12 months. Mean baseline interferon response gene score was not predictive of response to step-up DMARD therapy

The MBDA test score correlated positively with DAS28 at a single time point ( $r_s=0.58$ ,  $p<0.0001$ ) and the change correlated positively with corresponding changes in DAS28 ( $r_s=0.56$ ,  $p<0.0001$ ). The MBDA test categorised a higher proportion of participants with moderate and high disease activity than the DAS28; however, a notable proportion of high (58%) and moderate (59%) MBDA assessments were not associated with MSUS evidence of synovitis.

## **Conclusions**

MSUS and MBDA assessments of global disease activity identified active disease more frequently than corresponding DAS28 assessments. Compared to DAS28 driven therapy, MSUS driven step-up DMARD escalation was not associated with significantly better clinical outcomes but was associated with a higher rate of DAS44 remission at study completion. At present, there is no evidence to support the routine of MSUS to assess global RA disease activity; however, this position may change once the radiological and medium-long term outcomes are available. Peripheral blood gene expression analysis does not appear to contribute clinically useful additional information to the assessment of early RA. The MBDA test may provide an additional measure of disease activity; however, issues relating to specificity and its impact on clinical outcomes remain to be clarified.

# **1. Introduction**

Rheumatoid arthritis (RA) is a chronic, immune-mediated, inflammatory polyarthropathy, predominantly affecting the peripheral synovial joints, which can be associated with significant extra-articular and systemic comorbidities. If left either untreated, or inadequately treated, patients accumulate an increasing burden of erosive joint damage, progressive joint deformity, disability, socio-economic decline and premature mortality. Modern drug therapies and treatment regimens have significantly improved clinicians' ability to control the inflammatory process, so retarding, but not always preventing, structural and functional decline. Unfortunately, a significant subset of patients continue to experience persistently active disease and/or progressive joint damage. Recent advances in imaging technologies and understanding of RA pathogenesis offer new ways of 1. identifying persistent synovitis and 2. identifying those patients likely to be at an increased risk of either persistently active disease and/or progressive joint damage

## **1.1 Clinical Features**

Population studies performed in European and North America suggest that the prevalence of RA is between 0.5 and 1% (1,2). At onset, and during active phases of the disease, patients describe pain, stiffness and loss of function of the affected joints. Pain and stiffness are typically worse in the morning, or after resting, and are improved by repeated movements. Any synovial, diarthrodial joint can be involved, though typically patients present with a persistent, symmetrical peripheral polyarthritis affecting the metacarpophalangeal joints (MCPj) and/or proximal interphalangeal joints (PIPj) of the hands, the wrists and the metatarsophalangeal joints (MTPj) of the feet. (3,4). Larger joints (e.g elbows, shoulders, knees and hips) may also be affected but less commonly so. Patients may describe less specific systemic features such as lethargy, loss of appetite and weight loss. Clinical examination will identify the pattern of affected joints and confirm the presence of synovitis (the clinical expression of the inflammatory process). Synovitis is characterised by soft tissue swelling, with or without effusion, and tenderness related to the joint. It is often associated with loss of usual joint function. At presentation, joint deformities are unusual; however, patients with longstanding – or particularly aggressive – RA may exhibit subluxation and ulnar deviation of the wrists and MCPj, subluxation of the MTPj and characteristic hand deformities, such as swan necking and Boutonniere's deformity. As a systemic inflammatory condition, the clinical features of RA are not confined to the joints; a subset of patients may also develop extra-articular features, such as rheumatoid nodules, pulmonary fibrosis, pulmonary nodules, pleural and pericardial inflammation, pericardial effusions, splenomegaly and Felty's syndrome, vasculitis, mononeuritis multiplex and/or inflammatory eye lesions,

### **1.1.1 Diagnosis**

Traditionally, the diagnosis of RA was made on clinical and radiographic grounds by the identification of a symmetrical, peripheral inflammatory polyarthropathy affecting the small joints of the hands and feet and often in association with a positive, disease appropriate autoantibody (rheumatoid factor and/or anti-cyclic citrillunated peptides) and/or a characteristic extra-articular feature (such as rheumatoid nodules). However, this approach alone could lead to a delayed, or inappropriate diagnosis, in a significant subset of patients, since:

- I. The classical, clinical presentation can be emulated by other inflammatory conditions, such as psoriatic arthritis or polyarticular gout
- II. In the very early stages after symptom onset, a significant subset of patients will present with an asymmetric, inflammatory oligoarthritis which does not fit the classical clinical picture
- III. Rheumatoid factors (RF) and anti-cyclic citrillinated peptide antibodies (anti-CCP) can be detected in serum for several years before the onset of clinical disease (5-7)
- IV. RF has been identified in patients with a number of other rheumatological (e.g. Sjogren's syndrome, connective tissue disease, idiopathic inflammatory myopathies) and non-rheumatological disorders (e.g. chronic infection – notably subacute bacterial endocarditis, hepatitis B and hepatitis C – fibrotic pulmonary disorders, malignancy and primary biliary cirrhosis)(8,9) and also in healthy subjects
- V. RF and anti-CCP assays have only moderate sensitivity (48% and 54% respectively) (10); thus reliance on the presence of autoantibodies will misclassify a significant subset of patients

### **The role of disease classification criteria**

A number of classification criteria have been proposed to try and encourage uniformity of RA diagnosis in patients recruited to clinical trials. These criteria may estimate long-term prognosis based upon presenting features and may also facilitate early diagnosis by identifying those patients at an increased risk of progressing to rheumatoid arthritis from amongst all patients with undifferentiated inflammatory arthritis.

i. The 1987 American College of Rheumatology (ACR) Classification Criteria for RA (11) were developed to distinguish RA from other rheumatological conditions. They comprise:

1. early morning stiffness lasting more than one hour,
2. arthritis (soft tissue swelling) around three or more joint areas,
3. arthritis of hand joints – swelling in at least one area from wrist, MCP and PIP
4. symmetrical distribution of arthritis,
5. rheumatoid nodules,
6. presence of RF,
7. radiographic erosions, and/or periarticular osteopenia, affecting the hand and wrist joints on plain x-ray.

RA is confirmed by the presence of four or more of the criteria with criteria 1 – 4 having been present for at least 6 weeks. However, the criteria's ability to identify the early often non-erosive stages of RA is limited since they are prejudiced against patients with asymmetric, oligoarticular presentations and certain components (most notably radiographic erosions) are more commonly associated with longstanding RA. Whilst periarticular osteopenia and erosions are common in the early stages of RA (12) they are not specific for RA, having been described in association with other inflammatory arthritides (particularly psoriatic arthritis). Further, it is increasingly recognised that the 1987 ACR criteria are limited since they differentiate patients with established RA from patients with other inflammatory arthritides, an approach that runs contrary to the current consensus of early diagnosis and treatment (13)

ii. The 2010 Joint ACR and European League Against Rheumatism (EULAR) Classification Criteria for RA (13) were developed as a way of identifying those patients with undifferentiated inflammatory synovitis whose presenting features suggested that they were at a sufficiently high risk of developing either persistent inflammatory and/or erosive disease that they could be classified as having RA and commenced upon prompt immunomodulatory therapy (13). The classification algorithm requires the presence of at least one clinically swollen joint and then baseline clinical (symptom duration and distribution of clinical joint involvement), immunological (serology (RF and anti-CCP titres) and acute phase (CRP and ESR)) factors are scored depending upon their degree of involvement. A symptom duration of 6 weeks or greater gains additional weighting to allow differentiation from other, self limiting causes of inflammatory arthritis. When applied to the presenting features of a population of patients with undifferentiated inflammatory synovitis the 2010 ACR/EULAR Criteria classified a greater proportion of patients as RA, and allowed earlier introduction of disease modifying anti-rheumatic drug therapy (DMARD) than the 1987 ACR Criteria (14).

iii. The 'Visser Criteria' (15): prospective observation of 524 patients with early arthritis lead to the development of a statistical model which uses baseline presenting clinical (symptom duration, morning stiffness, arthritis in 3 or more joint groups, positive metatarsal squeeze test), immunological (IgM-RF and anti-CCP status) and radiological (presence of erosions on hand or foot xrays) factors to 1. determine the risk of a patient developing persistent arthritis (compared to self-limiting) and then 2. determine the risk of a patient with persistent arthritis developing erosive joint damage over 2 years follow-up. Receiver operating characteristic curve analysis demonstrated that this model's ability to discriminate persistent from self-limiting and erosive from non-erosive arthritis (i.e features supportive of the need for DMARD therapy) was significantly greater than the 1987 ACR Classification Criteria. The presumption being that persistent inflammatory arthritis, with a high risk of developing erosions, was highly likely to be RA

iv. The 'Leiden Prediction Rule' (16) comprises nine baseline demographic (age, sex) clinical (distribution of joint involvement, morning stiffness 100mm VAS, tender joint count, swollen joint count) and laboratory factors (C-reactive protein level, RF and anti-CCP status) that logistic regression analyses have shown to be independently predictive of the 1 year risk of developing RA in patients presenting with undifferentiated inflammatory arthritis. Higher scores are associated with a greater probability of developing RA. A subsequent validation exercise demonstrated that this prediction rule retained excellent discriminative ability to determine the likelihood of progressing to RA when applied to three geographically distinct European undifferentiated arthritis cohorts (17)

However, not all patients with undifferentiated inflammatory arthritis (UIA) will develop RA. Prospective observation of several inception cohorts has shown that approximately 60% of patients with UIA will spontaneously remit, approximately 30% will progress to RA and approximately 10% will develop an alternative inflammatory arthritis (18,19). The diagnostic and classification criteria described above serve as a means of differentiating those patients with very early RA, who may not yet exhibit classical clinical features but still require early DMARD therapy, from all other UA patients. In this way, patients with very early RA should not experience any

delay to the introduction of DMARD therapy and patients with UIA that is likely to remit will avoid the inherent risk of adverse effects associated with unnecessary therapy.

### **1.1.2 Clinical Course**

Following diagnosis, RA does not follow a uniform, or predictable clinical course. Spontaneous remission, without pharmacological intervention, is rare. Clinical course varies markedly between patients; some will experience a mild, virtually self-limiting condition whereas others may experience severe, rapidly progressive joint disease associated with significant systemic inflammation (4). Furthermore, the duration and magnitude of a positive response to an individual DMARD varies considerably between patients and, even with DMARD treatment, many patients will experience persistently active disease, progressive erosive joint damage and eventually functional decline (4,20). Three broad, long-term disease trajectories have been postulated (4,21):

1. Persistently active disease – persistent synovial and systemic inflammation causes progressive joint destruction, loss of functional ability and disability
2. Intermittently active disease – affects approximately 15-30% of patient. The level of synovial and systemic inflammation fluctuates and, with therapy, patients may experience prolonged periods of clinical remission. However, acute relapses may occur, involve new joint groups and may not be associated with an obvious precipitant. The risk of erosive progression is highest during an acute relapse
3. Prolonged clinical remission – is experienced by approximately 10% of patients. Patients may still experience occasional acute relapses; however, the long term risk of progressive joint damage and functional decline is low

### **The impact of disease activity and structural damage on functional capacity**

In many patients functional capacity often declines steadily over the disease course (22,23). However, at any given time point, the functional limitations imposed by RA are caused by a combination of the burden of inflammatory disease activity and the level of structural joint damage that has accumulated until that point. Over the long term course of RA, the balance, and relative importance, of these factors often shifts from the initial symptoms of active joint inflammation (i.e. disease activity) to the limitations caused by irreversible bone erosions and joint deformity (i.e. structural damage) (24).

Prospective observational studies have attempted to describe the impact, and relationship between, disease activity and structural damage on functional capacity in patients with newly diagnosed RA.

- Drossaers-Bakker et al described the functional outcomes in 138 females with newly diagnosed RA over 12 years of therapy (23). Functional capacity (Health Assessment Questionnaire (HAQ)) declined slowly over the follow-up period (baseline 0.63 (median); 12 years 0.87). Disease activity (44 joint Disease Activity Score (DAS44)) remained relatively stable (baseline = 2.9 (median); year 3 = 3.1; year 6 = 2.8; year 12 = 2.5) with a statistically

significant positive correlation between DAS44 and HAQ at each of the time points (Spearman's Correlation Coefficients: baseline = 0.68, year 3 = 0.51, year 6 = 0.52, year 12 = 0.61;  $p < 0.001$  each correlation). Structural damage (modified Sharp Score) declined throughout the study (total score: baseline = 0 (median); year 3 = 29; year 6 = 56; year 12 = 145) and was positively correlated with HAQ at each time point. A multivariate regression model demonstrated that DAS44 was the main determinant of functional capacity over the duration of the study and explained 51% of the variance in HAQ. Structural damage's contribution to the HAQ was of a lesser magnitude and, when added to the same multivariate regression model, explained an additional 12% of the variance in HAQ.

- Welsing et al observed 378 patients with early RA (< 1 years duration) for 9 years (24). Functional capacity (HAQ disability index (HAQ-DI)) declined over the follow-up period (baseline = 0.47 (median); 9 years = 0.63) at a decrement rate of 0.02 units per year. Disease activity levels (DAS44) improved initially and then remained constant. Correlation analyses demonstrated a statistically significant positive correlation between DAS44 and HAQ-DI at baseline, 3 and 6 years but not at 9 years (Pearson's Correlation Coefficients: 0.4, 0.4, 0.7 and -0.02 respectively). Structural damage (modified Sharp Score) worsened throughout the study (baseline = 11 (median); 9 years = 83.8), however did not exhibit a statistically significant positive correlation with HAQ-DI until 6 and 9 years of follow-up (Pearson's Correlation Coefficients: 0.15, 0.06, 0.75, 0.57). Furthermore, multiple linear regression analyses demonstrated that at 6 and 9 years follow-up structural damage modified, and down-played, the influence of DAS44 on HAQ.

Taken together, these results suggest that disease activity has the greatest impact on functional capacity in the early years after disease onset; whilst in the later years structural damage exerts a greater influence. However, disease activity and structural damage are not mutually exclusive. Clearly, there may be an inter-relationship between the total amount of inflammation a patient has been exposed to and the level of structural damage that they subsequently develop. Wick et al, demonstrated that joint destruction was a result of cumulative exposure to inflammation by constructing a mathematical model based upon the clinical and radiological outcomes of 76 patients with early RA (25). Cumulative inflammatory burden (area under the curve for DAS28) correlated positively with the observed joint destruction (modified Larsen Score) and was modified by a constant factor for each patient. Furthermore, whilst studies of plain radiographs (26-28) and MRI (29,30) have demonstrated that erosive structural damage is present in the very early stages of RA it is rare for bone marrow oedema (a marker for future erosions on MRI) to occur in the absence of synovitis (i.e active disease).

### **Predicting long-term prognosis in early rheumatoid arthritis**

Determining prognosis in RA is closely associated with trying to predict a patient's likelihood of response to disease modifying therapy and therefore their likely long-term inflammatory disease burden. Prognosis is not precisely defined, but most commonly refers to a continuum of disease outcomes beginning with the likelihood of a patient responding to disease modifying therapy, developing progressive joint erosions, joint deformity loss of function and eventually long-term

disability (Figure 1).

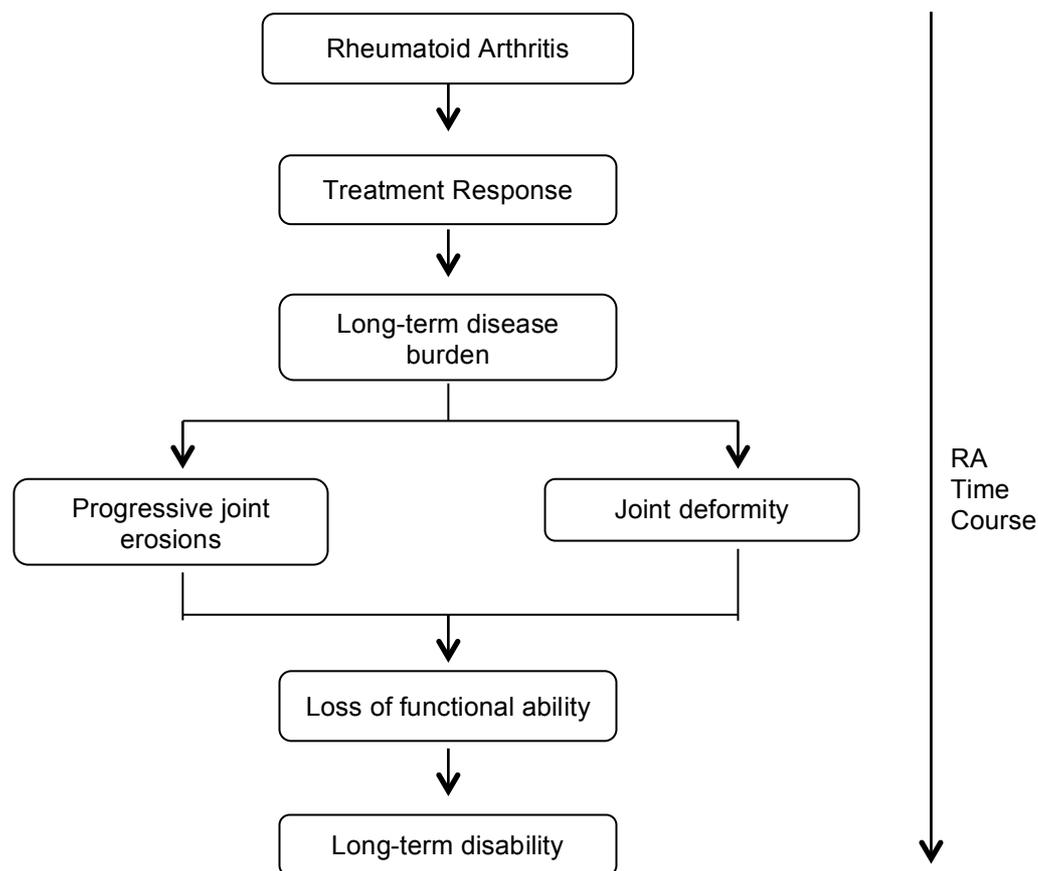
Identifying patients with poor prognostic features (i.e. those at the highest risk of persistently active and progressive disease) may allow tailoring of therapy to individual patient's needs. Based on presenting features, patients considered to have a comparatively poor long-term prognosis could be identified at outset and 'triaged' to receive a more aggressive initial treatment regimen compared to those with a more favourable prognostic profile.

A number of demographic and disease related factors have been shown to have prognostic properties and can be divided into those that are fixed and those that are potentially modifiable (31,32):

### ***Fixed predictors of prognosis***

*Age and Gender* – It is well established that there is a higher incidence and prevalence of RA amongst women than men (33-35). A large, cross-sectional cohort study in 6004 patients has demonstrated that women consistently exhibit higher scores for disease activity (ACR Core set: DAS28, 28 tender and swollen joint counts, ESR, global health VAS, pain VAS and physician global estimate), lower remission rates, worse functional ability (HAQ) and have a higher prevalence of erosions compared to men (36). Furthermore, several longitudinal studies of patients with newly diagnosed, and untreated, RA have demonstrated that women tend to experience persistently higher overall measures of disease activity, lesser rates of remission and greater degrees of functional decline even though disease characteristics at presentation are similar to men (37-40).

Considering mortality, younger age at symptom onset (< 55 years) has been associated with a higher risk of death from cardiovascular disease (41,42). However, the influence of age at onset on RA outcomes is less clear and the evidence is often conflicting. Prospective follow-up (median 3.6 years) of 400 patients with newly diagnosed RA demonstrated that those with late onset disease (age > 65 years) experienced similar changes in disease activity, radiographic progression and functional ability, and higher remission rates, compared to those with early onset disease (age < 65 years) (43). However, conversely, step-wise regression analysis of a six year study of 332 patients with early RA showed that older age at presentation was predictive of higher disease activity measures, higher rates of radiographic progression and worse functional ability over the follow-up period (44). Camacho et al investigated the inter-relationship between age at presentation and gender on RA prognosis by examining the disease activity and functional ability outcomes of 3666 patients with recent onset inflammatory polyarthritis (45). At presentation women of all ages had similar levels of functional impairment and, overall, women generally had higher levels of functional impairment than men at any given follow-up time point. However, beyond 5 years of follow-up women with very late-onset disease (age > 75 years at presentation) experienced a more rapid acceleration in functional decline than those presenting at an earlier age. These results suggest that the impact of sex on RA disease activity is evident at presentation; however, the impact of age at presentation on functional ability, and therefore prognosis, may not become evident until much later in the follow-up period



**Figure 1** – Prognostic continuum in rheumatoid arthritis

*Auto-antibody Status* – Rheumatoid factors (RhF) and anti-citrullinated protein antibodies (ACPA) are important diagnostic indicators for RA and feature prominently in the 2010 ACR/EULAR RA Classification Criteria (13,46,47). RhF seropositivity has been associated with an increased rate of extra-articular manifestations (e.g. rheumatoid nodules, rheumatoid vasculitis) (48), increased mortality (49,50) and faster rates of destructive radiographic progression over 3–10 years follow-up (47,51). Equally, possession of ACPA antibodies has consistently been shown to be associated with higher measures of disease activity and rates of radiographic damage progression over time. (1,8,51-53). Whilst presenting clinical disease activity measures and radiographic findings were similar for patients with ACPA-positive and ACPA-negative RA, van der Helm-van Mil et al demonstrated, that on longitudinal follow-up, ACPA-positive patients exhibited persistently higher swollen joint counts (a surrogate for clinical synovitis) and significantly higher scores for radiological damage (total Sharp-van der Heijde Score;  $p < 0.001$ ) (52). Similarly, Ronnelid et al showed that, over 5 years follow-up, ACPA-positive patients had experienced persistently, statistically significant, higher levels of disease activity (ESR, swollen joint count, DAS28) and higher rates of radiological progression (change in Larsen Score (mean): 9.7 vs 6.9;  $p = 0.01$ ) compared to ACPA-negative patients who had otherwise similar presenting features and received a similar intensity of immunomodulatory therapy (1).

*Radiographic Evidence of Structural Damage* – Patients with evidence of structural joint damage on presenting plain x-rays of hands and feet are at an increased risk of accumulating further damage in the future (3,5,54). Post-hoc analysis of 870 patients recruited to the ASPIRE study

found that patients with evidence of structural damage at baseline were at a greater risk of developing further damage after 1 year than those with no structural damage (5). Furthermore, patients were more likely to develop worsening of whichever radiographic finding (i.e. joint space narrowing or erosions) was particularly prevalent on the baseline radiographs.

*Genetic Factors* – Multiple studies have identified specific genes associated with RA susceptibility. There has been far less research into whether or not specific genes are associated with treatment outcome and prognosis. A meta-analysis of 29 studies did demonstrate that, in most populations, possession of the shared epitope (HLA-DRB1) was associated with a significantly higher risk of developing plain x-ray erosions (OR 2.0; 95% CI 1.8-2.2) (8). Furthermore, several studies have shown that possession of specific genetic polymorphisms is associated with the likelihood of a positive treatment response (and thereby indirectly imply prognosis) to either methotrexate (55) or anti-TNF alpha blocking therapy (56,57). Pharmacogenetic analysis of the BeST Study has shown that, in early RA, carriage of the AMPD1 34T, ATIC 347CC, and/or ITPA 94CC alleles is associated with a good clinical response (achieving DAS44 < 2.4 after 6 months therapy) to methotrexate monotherapy which is particularly evident when all three alleles are present (OR 27.8; 95% CI 3.2-250) (55). In established RA, possession of specific single nucleotide polymorphisms (SNPs) within specific candidate genes, relating to the Toll-like receptor and NFkB signalling pathways, have shown a positive association with the absolute change in clinical disease activity (CHUK, IKBKB, MyD88, NFkB1A, TLR-2, TLR-4) and likelihood of achieving a moderate-good EULAR response (CHUK, IKBKB, MyD88, TLR-2, IRAK3, NFkB-2, NFkBIB, PTGS2, TLR10/1/6), following commencement of anti-TNF $\alpha$  blocking therapy (56). Furthermore, in a similar population, possession of SNP variants in two RA susceptibility genes (AFF3, CD226) has also been shown to have a statistically significant association with the observed clinical response to anti-TNF $\alpha$  blocking therapy (57); thus suggesting that particular susceptibility alleles might also have implications for treatment response and feasibly prognosis

### ***Potentially modifiable predictors of prognosis***

Modifiable predictors of prognosis relate to either those external factors which have been strongly linked to RA pathogenesis, and can potentially be removed through lifestyle / environmental adjustment (most notably smoking), or those disease specific factors that therapeutic intervention aims to influence

*Tobacco Smoking* – Tobacco smoking has implications for both RA susceptibility and long-term prognosis following clinical disease onset. Compared to never-smokers, those who smoke more than 25 cigarettes per day are at 32% increased risk of developing RA (OR 1.32; 95% CI 1.19-1.46) and a 44% increased risk of developing RhF-positive RA (OR 1.44; 95% CI 1.23-1.65) (58). In the same analysis, longer periods of cigarette smoking exposure exerted a greater influence on the likelihood of developing both RA and RhF-positive RA than the quantity of cigarettes being smoked. Furthermore, there is a strong gene-environmental interaction evident in cigarette smokers who may already be genetically predisposed to develop RA. Individuals who are

homozygous for the HLA-DRB1 shared epitope and are cigarette smokers are at a significantly increased risk of developing ACPA-positive RA compared to shared epitope positive non-smokers (OR 17.8; CI 10.8-29.4)(59,60).

Cohort studies have also demonstrated that cigarette smoking increases an individual's risk of developing RA that exhibits other poor prognostic features. Once RA is established, cigarette smoking has been associated with an increased risk of extra-articular disease (most notably rheumatoid nodules and pulmonary complications), lower measures of functional ability and a higher burden of radiological joint damage (61,62). Furthermore, cigarette smoking may influence longer-term prognosis by modulating a patient's treatment response to both conventional DMARD and biologic therapy. A prospective study of 225 patients with early RA (<2 years duration) demonstrated that cigarette smoking was the only factor significantly associated with the likelihood of a patient not achieving ACR50 response after 6 months of combination DMARD therapy (OR 3.91; 95% CI 1.41–10.81) (21). Additionally, of 1430 patients with early arthritis, those who smoked were significantly less likely to achieve a EULAR good response after 3 months of therapy with either methotrexate (n = 873; 27% vs 36%, p = 0.05) or anti-TNF $\alpha$  blocking therapy (22). Taken together, all of these results suggest cigarette smoking is likely to be a poor prognostic factor since it increases the likelihood of developing RA (particularly in individuals who are already at an increased genetic risk), is associated with an increased risk of developing poor prognostic features and is associated with lesser treatment responses

*Time to Initiation of DMARD Therapy* – Following onset of clinical disease a therapeutic window has been proposed when the emerging inflammatory process is considered most amenable to intervention and therefore most likely to respond positively to therapy. Certainly, delays in commencement of DMARD therapy have consistently been associated with lesser treatment responses and poorer long-term clinical and functional outcomes (7,63). Pooling outcome data from 14 randomised clinical trials of DMARD therapy in RA (1435 patients in total) demonstrated that there is an incremental decline in treatment response the longer after symptom onset that therapy is initiated (mean ACR response rates: <1 year duration 53%, 1-2 years 43%, 2-5 years 44%, 5-10 years 38%, >10 years 35%; p = 0.001) (64). Similarly, longer disease durations prior to commencing DMARD therapy have been associated with significantly lesser chances of achieving clinical remission (65,66) and lesser improvements in functional ability during therapy (67). Whilst delays in commencement of DMARD therapy adversely affect short-term therapeutic response, they also have important prognostic implications since they are often eventually associated with worse long-term radiological outcomes. Meta-analysis of 12 randomised clinical trials of DMARD therapy in early RA has demonstrated that early initiation of DMARD therapy resulted in a 33% lesser rate of radiographic progression compared to delayed therapy (mean delay = 9 months) (26).

*Disease activity at presentation* – Long-term cohort studies have demonstrated that patients who present with low levels of inflammatory disease activity and lesser degrees of functional impairment have comparatively better prognoses than those who present with high levels of inflammatory disease activity. Gossec et al prospectively followed 191 patients with early

rheumatoid arthritis for 5 years. Those patients who presented with lower clinical measures of disease activity (DAS44 < 4; RAI < 17), lesser degrees of functional impairment (HAQ < 1.25) and lower laboratory measures of inflammation (CRP < 14.5mg/l) were significantly more likely to have achieved clinical remission by 3 years follow-up and sustained remission after 5 years follow-up (29). By inference, those patients who present with high measures of disease activity, functional impairment and/or acute phase response would be expected to have a comparatively poorer treatment response and thus prognosis. To some extent this has been borne out. Follow-up of 191 early RA patients has shown that high baseline measures of CRP and ESR (surrogates for disease activity) were independently predictive of the degree of functional impairment (HAQ) evident after 5 years of treatment (31).

*Longitudinal disease activity* – As a chronic inflammatory disease, RA is subject to fluctuations in overall disease activity. Indeed, treated patients remain at risk of acute flares following either loss of treatment effect, and/or exposure to an external precipitating factor. Even taking into account the factors described in the preceding sections, long-term prognosis will still depend heavily upon how much persistent inflammatory disease activity that a patient is exposed to over time. Therefore, cumulative inflammatory disease burden, a reflection of overall treatment response which does not relate to a specific treatment regimen, will be an important determinant of prognosis but may not be immediately evident based upon presenting features alone. Persistent elevation of acute phase reactants (ESR and CRP), thereby suggesting persistent inflammatory disease activity, has been associated with progressive accumulation of erosive radiographic damage (33,35). Furthermore, persistent elevation of clinical disease activity measures has also been associated with functional decline and radiographic progression (37,68). Using a cohort of 194 early RA patients treated with step-up DMARD therapy, Conaghan et al demonstrated that a higher proportion of patients with persistently high DAS28 (>5.1) and persistently moderate DAS28 (>3.2 and < 5.1) experienced declines in functional ability (HAQ) over 12 months follow-up compared to those patients who had persistently low DAS 28 (<3.2) (percentage with deterioration in HAQ; high DAS28 46.7, moderate DAS28 21.4, low DAS28 10.9) (37). Similarly, Salaffi et al have demonstrated that over three years follow-up persistent elevation of disease activity (DAS28-CRP) was predictive of progressive radiographic damage (41). At each time point, compared to patients without radiographic progression, those patients who exhibited the greatest rate of radiographic progression consistently exhibited significantly higher levels of inflammatory disease activity and a greater cumulative inflammatory disease burden (AUC: DAS28-CRP).

### ***Biological markers of prognosis***

Multiple studies have identified biological markers, often associated with abnormal immunological and inflammatory processes, that are either predictive of treatment response and/or long term prognosis. However, few of these markers are used in routine clinical practice and the majority require additional validation exercises. The preceding sections have described the potential prognostic properties of commonly measured immunological (RhF and ACPA) and acute phase response (CRP and ESR) markers. The following sections will describe the potential prognostic properties of several investigational biological markers. In many cases, individual markers have

been identified through exploratory studies attempting to identify predictive markers of response to specific biologic agents and their prognostic properties in early RA must be inferred (69).

*Immunological Factors* – several additional antibodies have demonstrated diagnostic properties which allow either clear differentiation of RA from other causes of inflammatory arthritis or suggest RA in patients who are seronegative for RhF and/or ACPA (70). Amongst these a number have also demonstrated some relationship to treatment response:

i. Anti-epidermal filaggrin antibodies (anti-keratin and anti-perinuclear factor) have been associated with persistent disease activity, and therefore treatment resistance, in early RA but do not appear to be associated with subsequent radiographic progression (46,47).

ii. Anti-mutated citrullinated vimentin antibodies (anti-MCV; anti-Sa) may be linked to inflammatory disease activity since anti-MCV titres have been shown to correlate strongly with disease activity measures (DAS28;  $r = 0.5334$ ;  $p = 0.0003$ ) and allowed stratification of patients into groups based upon disease activity (71). Furthermore, patients with early RA who express anti-MCV antibodies have been shown to experience a significantly lesser treatment response and a greater overall inflammatory disease burden (DAS28 AUC) compared to RA patients who are anti-MCV negative (49).

iii. Increased levels of IgG lacking galactose (termed Gal 0 glycoforms) have been associated with RA and have been correlated to disease activity. In female RA patients, who subsequently become pregnant, Gal 0 Glycoform levels are elevated in the pre-partum period, fall with pregnancy associated disease remission and re-increase in the post-partum period (72)

*Genetic Factors* – particular characteristics of an individual's genotype may influence their risk of developing RA and may also bear upon its long term severity once joint disease has manifested:

i. Shared Epitope – Possession of specific HLA-DR allele variants (particularly HLA DR4) is associated with an increased risk of developing RA in the future. Furthermore, possession of the shared epitope has prognostic has been associated with an increased risk of developing joint erosions (8) and of possessing ACPA antibodies (53); both of which have independently been shown to be poor prognostic markers

ii. Matrix Metalloproteinase Genotype – The matrix metalloproteinases are a group of enzymes that contribute to erosion formation through the degradation of collagen and cartilage. Possession of a specific polymorphism in the matrix metalloproteinase-3 gene (MMP3 6A/6A) appears to have prognostic implications since in a single longitudinal study of 103 early RA patients it was positively associated with a greater degree of erosive joint damage at presentation and a significantly higher rate of radiographic progression over 4 years follow-up (73).

iii. Interleukin-10 Promotor Genotype – A single prospective study in 283 RA patients demonstrated that specific polymorphisms of the interleukin-10 promotor gene, may have

prognostic implications. Patients possessing alleles which coded for high levels of IL-10 production (-2849 AG/GG) had higher titres of RA associated autoantibodies, and experienced greater degrees of radiographic progression over 2 years follow up, compared to those who possessed alleles coding for low levels of IL-10 production (-2849 AA) (54).

*Acute Phase Reactants* – the potential prognostic role for commonly used measures of the acute phase response, CRP and ESR, has been discussed in the preceding section 2.1.2.2.2. Interleukin-6 (IL-6) is released by inflamed synovial membrane into blood and has several systemic effects through; regulation of platelet production, development of the anaemia of chronic disease and stimulation of the liver to synthesise acute phase proteins. However, whilst plasma IL-6 levels correlate positively with measures of the acute phase response, it is unlikely that IL-6 has any prognostic properties since no apparent relationship between plasma IL-6 levels and the rate of radiographic progression has been demonstrated (74).

*Tissue Specific Markers* – Since inflamed synovium can damage several different tissue layers of an affected joint it is feasible that the expression levels of markers related to metabolism / degradation within each of these tissue compartments might also have prognostic properties. However, their use in routine practice remains unclear since measurement of the marker often requires percutaneous biopsy procedures to harvest the target tissue

i. *Synovial Membrane Markers*: hyaluronan is a glycosaminoglycan released by inflamed synovial membrane that can leak into the circulation and be measured in high levels in the serum of RA patients. Hyaluronan may be a marker of on going joint damage since, in a prospective study of 40 RA patients, serum levels correlated positively with radiographic damage scores at presentation and remained elevated in patients who demonstrated progressive radiological joint destruction over 12 years follow-up (75). Matrix metalloproteinases (MMP) are released by inflamed synovium and contribute to joint damage by mediating cartilage degradation. Therefore, it is feasible that their persistent activity (or persistent elevation) could have implications for long-term prognosis. In 98 patients with untreated RA, baseline levels of MMP-1 and MMP-3 correlated positively with CRP (as a measure of baseline disease activity) and with the rate of change of radiographic damage (Larsen Index) and functional decline (HAQ) after 12 months follow-up (76). Furthermore, a logistic regression analysis identified baseline MMP-3 levels as the strongest predictor of developing radiographic damage in patients who initially had non-erosive x-rays.

ii. *Cartilage-Specific Markers*: elevated levels of markers of cartilage metabolism have been shown to have both negative and positive prognostic implications. Using hip joint destruction as a marker of radiographic progression, all patients with rapid radiographic progression exhibited elevated levels of cartilage oligomeric matrix protein (COMP). Patients with slow radiographic progression had significantly lower levels of COMP but higher levels of chondroitin sulphate epitope 846, a marker of cartilage aggrecan synthesis (77). Elevated baseline urinary excretion of crosslinked c-terminal peptides from type II collagen (CTX-II) and degradation products of the

helical region of type II collagen (Helix-II), have been shown to correlate positively with changes in radiographic damage over 12 months follow-up (CTX-II  $r = 0.3$ ,  $p = 0.007$ ; Helix-II  $r = 0.22$ ,  $p = 0.05$ ) (78). Furthermore, patients who exhibited elevated levels of both markers experienced higher rates of radiographic progression than those who exhibited elevation of either one or neither marker.

iii. Bone-Specific Markers: several markers, specific to bone metabolism, have been linked to the development of progressive joint destruction in RA. Synovial fluid, but not serum, levels of bone sialoprotein, a protein released by osteoblasts in juxta-articular bone, have correlated positively with increasing degrees of joint destruction on knee radiographs (79). However, this finding is not confined to RA and was also demonstrated in patients with osteoarthritis. Prospective studies have shown that elevated serum and urinary levels of cross-linked carboxyterminal telopeptides of type I collagen (ICTP) both correlate positively with measures of RA disease activity and appear predictive of future radiographic progression (46,80). Furthermore, persistent elevation of serum ICTP levels despite 6 months of DMARD treatment has been associated with increased rates of radiographic progression compared to those patients whose serum levels fall with treatment (81).

iv. Vascular Markers – Whilst RA predominantly causes pathological changes within the synovial membrane of affected joints it is also a systemic disease and specific, pathological changes have been frequently described within the systemic vascular bed. Hence, markers of RA-related vascular inflammation might also have prognostic implications for the overall disease process. Vascular endothelial growth factor (VEGF) is expressed at increased levels in serum and synovial fluid in RA patients. In early RA, serum VEGF levels have been shown to correlate positively with clinical measures of disease activity (swollen and tender joint counts) and reflect treatment response since they decrease significantly in patients who achieve moderate-good EULAR response rates after DMARD therapy (59). Furthermore, baseline serum VEGF levels positively correlated with, and therefore may be predictive of, subsequent changes in radiological damage scores (Spearman's  $r = 0.579$ ,  $p = 0.004$ )

To date studies attempting to identify biological markers of prognosis have tended to focus on single candidates or families of markers related to activity within a single inflammatory process or tissue compartment. However, for a disease with such a widely heterogeneous phenotype and clinical course as RA, it is quite possible that relying upon a single marker to provide prognostic information for all patients will prove inadequate. Alternatively, combining and comparing the expression of several biological markers, with known prognostic properties and representing the different genetic and cellular layers of the inflammatory disease process, might give a more accurate, and nuanced, indication of an individual patient's long term disease course and likelihood of responding to therapy.

- Proteomic analysis of serum from 44 patients with established RA identified a panel of proteins (IL-6, IL-2, oncostatin M, macrophage colony-stimulating factor (M-CSF), tumour necrosis factor receptor superfamily member 9 (TNFRSF9), CCL23, transforming growth

factor-alpha (TGF-alpha), CXCL13) which correlated positively with traditional disease activity measures and were differentially expressed between patients with RA judged to have either active or inactive disease (82). Furthermore, multivariate analysis created a statistical model comprising 5 markers (CXCL13, CCL23, TGF-alpha, TNFRSF9, M-CSF) which accurately predicted the disease activity level (DAS28) at the time of testing.

- In a longitudinal analysis of 118 patients with early RA receiving DMARD therapy, Young-Sim et al investigated the ability of baseline traditional clinical and laboratory disease activity measures in combination with serum and urinary levels of candidate markers for synovial inflammation and cartilage turnover to predict progression of radiographic damage after 2 years follow-up (83). Multivariate logistic analysis identified elevated baseline levels of serum MMP3 and urinary CTX-II as being the only two factors which were independently predictive of subsequent radiographic progression (PPV 62.1 and 57.7 respectively).

### **1.1.3 The Pre-symptomatic Stages of Rheumatoid Arthritis**

It is not possible to make the diagnosis of RA until patients develop clinical signs and symptoms that suggest inflammatory joint disease (i.e. clinical synovitis). In many cases, the diagnosis is made on the basis of clinical features 'fitting' the typical description of RA and the presence of particular autoantibodies lends weight to the clinically suspected diagnosis. In cases where patients have evidence of inflammatory joint disease but do not 'fit' the typical description of RA, the presence of either RhF and/or ACPA antibodies can be used to estimate the likelihood that the patient is displaying an atypical or early presentation of RA or to estimate their risk of progressing to RA in the future. Several observational cohort studies have clearly demonstrated that subjects who eventually develop RA display evidence of abnormal immune activation and auto-antibody production for several years before the development of symptomatic joint disease (6,7,22,84-86). There is also a dynamic element to the autoantibody production; prior to the onset of symptomatic joint disease subjects exhibit a sharp increase in the overall titre of serum ACPA antibodies (7) and a marked expansion in the number of citrillunated epitopes that are recognised by ACPA antibodies (epitope spreading) (87-89). In addition to the presence of ACPA antibodies a number of other factors have also been suggested to increase the risk of a an at-risk subject eventually developing RA:

*Tobacco Smoking* – Tobacco smoke exposure is an important component of a complex gene-environmental interaction whereby subjects who smoke are at increased risk of developing RA, particularly if they already possess a genetic predisposition (60,90). Epidemiological studies of monozygotic twins and large prospective cohorts had previously recognised that subjects who smoked were more likely to develop RA than those who didn't smoke (58,91), with the stronger determinant being duration of smoking, rather than volume. Several linked studies have demonstrated that the risk is particularly prevalent in smokers who also possess a genetic predisposition to the development of RA, such as the shared epitope of HLA-DRB1 or polymorphisms of the PTPN22 gene (60,90,92,93). Indeed, current smokers, who are homozygous for the HLA-DRB1 shared epitope have been found to have a 15-23 times increased

risk of developing RA compared to non-smokers who do not possess the shared epitope gene. As previously discussed, tobacco smoke exposure has consistently been shown to increase the risk of developing seropositive (Rhf and/or anti-CCP antibody) RA. Recently it has been postulated that the influence of tobacco smoke on development of RA is mediated through abnormal immune activation, and particularly the development of autoreactivity to citrillinated peptides, at mucosal surfaces; such as the gums (94-96) or bronchial epithelium (97,98)

*Infection* – The initial trigger for autoreactivity and ACPA antibody production prior to the development of clinical RA is incompletely understood. There is increasing evidence that the initial immune event may be independent of the synovium, particularly since RhF and ACPA antibodies are frequently detected in the absence of synovial inflammation (99). Further, there is an emerging consensus that infective and immune episodes that occur at mucosal sites (such as the gum and/or respiratory epithelium) may play a role in the initial synthesis of ACPA antibodies (94,98,100). Recently, it has been recognised that the presence of certain bacteria (particularly porphyromonas gingivalis) within the oral biofilm, and the development of periodontitis, is a risk factor for both the production of ACPA antibodies (94,96) and the future development of autoantibody positive RA (101)

*Hormonal Factors* – The potential influence of hormonal factors on the subsequent development of RA is well recognised. Women are more frequently affected than men (102), the peak incidence occurs after the menopause and periods of hormonal flux, such as the post-partum or peri-menopausal periods, are frequently associated with the development of RA (103,104). There is conflicting evidence about the link between hormonal exposure and the development of auto-antibodies in the pre-clinical stages of RA. Some observational studies have suggested that increased hormonal exposure (such as early menarche or oral contraceptive pill use) may increase the risk of developing anti-CCP antibodies (105), whereas alternative studies have suggested that oral contraceptive pill use was protective against the development of rheumatoid factors (106). There is little published research describing the link between hormonal factors and future risk of RA in asymptomatic subjects who express RA associated auto-antibodies

*Obesity* – Obesity is often considered a state of chronic low grade inflammation and being obese has been associated with the development of several different inflammatory conditions. Large scale epidemiological studies have observed that there may be a link between body mass index and risk of developing RA because there were higher proportions of obese patients within the RA group than the unaffected control groups (107,108). However, interestingly, recent studies of the presenting characteristics of new RA patients have suggested that patients with auto-antibody negative RA have significantly higher body mass indices than those with auto-antibody positive RA (105,109). The risk association between body mass index and later development of RA in asymptomatic, anti-CCP positive individuals has not yet been described

*Alcohol Intake* – Several, independent case-control studies have demonstrated that the level of alcohol consumption in unaffected control subjects is statistically higher than in incident RA patients; the inference being that alcohol may have a protective effect on the development of

RA(105,107,110). Further, a retrospective analysis conducted using blood samples donated to the Nurses Health Study has suggested that daily alcohol consumption was associated with lower levels of pro-inflammatory markers (IL-6, sRNFRII), though not anti-CCP titres, in the asymptomatic stages before the onset of clinical RA

## **1.2 Current Approaches to the Management of Early Rheumatoid Arthritis**

### **1.2.1 Core Principles**

It is now widely accepted that DMARD therapy should be commenced as soon as possible after the clinical diagnosis of rheumatoid arthritis can be made in order to retard the development of early, irreversible joint damage and long term functional impairment (111,112). Several core principles underpin most commonly accepted RA treatment strategies

#### ***Early Commencement of DMARD Therapy***

In the early stages of RA there appears to be a therapeutic window when the nascent inflammatory process is most likely to respond to therapy (the proposed window of opportunity) (113). The timing and duration of any therapeutic window will likely vary between individual patients. Both relatively short delays in the commencement of DMARD therapy, and presentations with longer symptom durations, have consistently been associated with lesser treatment responses and poorer functional outcomes. Importantly, short-term delays in therapy have long-term adverse consequences and, similarly, early control of inflammatory disease activity appears to have long-term benefits. However, despite several studies reporting similar results the observation remains an association, rather than a causal relationship. A number of the older studies may have been confounded by either not correcting for disease severity at presentation and/or the use of low-intensity DMARD regimens that are contrary to current treat-to-target principles.

- The FINRACo trial (FINnish Rheumatoid Arthritis Combination therapy trial), and its follow-up studies, assessed the short, medium and long-term impact of different initial DMARD regimens in 195 patients with newly diagnosed RA (65,114,115). Patients were randomised to receive either sequential DMARD monotherapy (initially sulfasalazine) or combination DMARD therapy (methotrexate, sulfasalazine, hydroxychloroquine and prednisolone) with both groups aiming for clinical remission (ACR definition). After 2 years follow-up, in the sequential DMARD monotherapy arm, significantly fewer patients with a longer symptom duration prior to commencing treatment had achieved remission than those with a shorter symptom duration (ACR remission rate; symptom duration < 4 months = 35%, symptom duration > 4 months = 11%;  $p = 0.021$ ) (65). Symptom duration prior to DMARD commencement did not affect the likelihood of achieving remission in the combination therapy arm suggesting that the adverse impact of delayed therapy might potentially be attenuated by using a more aggressive treatment regimen at outset

- Green et al investigated the apparent relationship between presenting disease characteristics and the likelihood of persistent disease activity in 63 patients with undifferentiated inflammatory arthritis (51% fulfilled 1987 ACR RA classification criteria) (66). A regression analysis, incorporating multiple clinical and laboratory disease activity variables, demonstrated that disease duration prior to receiving treatment was the only factor independently associated with the likelihood of patients experiencing persistent disease activity after 6 months follow-up (median symptom duration (IQR): persistent disease activity group = 20 weeks (12-32), disease remission = 10 weeks (8-20),  $p < 0.05$ )
- Nell et al performed a case-control study comparing clinical and radiological outcomes in patients with early (< 3 months) and late (3 – 12 months) presentations of RA (116). At all time points over a 36 month follow-up period, patients who presented with early symptoms experienced significantly greater treatment responses ( $\Delta$ DAS28, ACR20/50/70 responses) lesser measures of disease activity (DAS28), greater improvements in functional ability (HAQ) and lower radiographic damage scores (Larsen Index). Almost identical outcomes were observed when a second early RA cohort were followed over a similar period
- Analysis of patients referred to the Leiden Early Arthritis Clinic has demonstrated that those who had longer symptom durations prior to commencement of DMARD therapy (ie delayed therapy), were less likely to achieve remission, and were more likely to experience radiographic deterioration than those with shorter symptom durations (117). Of 598 patients with RA, 412 (86%) were assessed after 12 weeks of symptom onset and were considered to have received delayed DMARD therapy. Over a 6 year follow-up period, those patients who received delayed DMARD therapy exhibited significantly higher scores for radiographic damage at all time points, regardless of the favoured DMARD regimen. Furthermore, patients in the delayed therapy group were significantly less likely to achieve sustained, drug-free remission (hazard ratio: 1.8 [95% CI 1.17-3.0],  $p = 0.009$ ). Older age at onset, gradual symptom onset, small joint involvement, presence of RhF and anti-CCP antibodies and low CRP levels were each independently associated with delay in review by a rheumatologist; however, their individual relationship to observed outcomes was not described

### ***Early Tight Control of Disease Activity***

If cumulative total exposure to inflammatory disease activity (ie the inflammatory burden) is associated with worse clinical and radiological outcomes it is reasonable to assume that an RA patient's longer-term prognosis can be positively influenced by DMARD treatment regimens which aim to minimise overall exposure to inflammation by being intolerant of persistently active inflammatory disease. To some extent this presupposition has been borne out by several strategic RA treatment studies which used persistent evidence of inflammatory disease activity to trigger escalation in a patient's DMARD therapy.

- The Tight Control of Rheumatoid Arthritis Trial (TICORA) demonstrated that newly diagnosed RA patients who underwent regular review (monthly), formal quantification of global disease activity (DAS44) and whose initial step-up DMARD treatment strategy was escalated aggressively if disease activity assessment exceeded a predefined threshold (DAS44 > 2.4), experienced significantly greater clinical improvements and less radiological progression compared to similar patients who underwent less frequent review and received a less aggressive DMARD treatment strategy that was guided by clinical findings rather than a defined disease activity target (118). 111 patients with newly diagnosed RA were randomised to either an intensive or routine treatment and follow-up strategy. At all follow-up time points, patients in the intensive group exhibited significantly lower disease activity scores and an overall greater improvement in all measures of disease activity, quality of life and functional ability. Furthermore, after 18 months follow-up patients in the intensive strategy group had experienced significantly lesser changes in radiographic erosion and total Sharp scores and lesser (not statistically significant) changes in joint space narrowing scores. It is worth noting that whilst presenting clinical and demographic features for both groups were similar the mean disease duration was 19 months; therefore, the relatively late presentation and commencement of therapy for all patients may actually have had a negative impact on the impressive outcome results.
- The Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA) study (119) had a similar design to TICORA though utilised percentage change in disease activity rather than a composite disease activity measure. Two hundred and ninety-nine patients with early RA (symptom duration < 2 years) were randomly assigned to either an intensive or conventional treatment strategy group. Patients in the intensive strategy group were reviewed monthly and treatment escalation decisions were based upon the output of a computer based decision making tool which analysed the change in clinical and laboratory disease measures. Patients in the conventional strategy group were reviewed monthly and DMARD escalation decisions were at the discretion of the treating clinician. Over a 2 year follow-up period, patients in the intensive strategy group experienced faster falls in clinical and laboratory disease activity measures, a higher chance of attaining clinical remission (50% vs 37%;  $p = 0.029$ ) and significantly longer periods of remission (mean duration: 11.6 vs 9.1 months;  $p=0.025$ ). Of the patients who did demonstrate radiographic change, progression rates tended to be higher in the conventional group.

Taken together the results of the TICORA and CAMERA studies demonstrate that patients' early response to DMARD therapy can be significantly improved through frequent reviews, formalised assessment of global disease activity and aggressive escalation of DMARD therapy in the presence of persistent disease activity. Importantly, these benefits were observed in patients who would now be considered to have presented relatively late (i.e. outwith the proposed window of opportunity).

Follow-up analyses of the FINRaCo trial have demonstrated that good short-term treatment responses have a positive medium-longer term impact. After 5 years follow-up and unrestricted

DMARD therapy, disease activity measures in the initial combination therapy and sequential therapy groups were similar, however patients in the combination therapy group demonstrated lower radiographic damage scores (median Larsen Index: 11 vs 24;  $p = 0.001$ ) and a significantly lesser accrual of radiographic damage (114). After 11 years follow-up, the benefits from initial early aggressive therapy remained evident. Functional ability scores were similar between treatment groups (mean HAQ: combination therapy = 0.34, sequential therapy = 0.38); however, the combination therapy group demonstrated significantly higher remission rates (ACR remission rate: combination therapy = 37%, sequential therapy 19%), a significantly higher proportion had achieved minimal disease activity (combination therapy = 3%, sequential therapy = 43%,  $p = 0.016$ ) and had a greater overall chance of ever attaining ACR remission at any time point endpoint (115)

### ***Predefined Disease Activity Target***

The treatment regimens employed by the FINRaCO, TICORA and CAMERA studies all incorporated a formalised assessment of global disease activity and a threshold measure above which escalations in DMARD therapy would be considered. The FINRaCO study threshold was less than 50% improvement in any two of three criteria (swollen joint count, tender joint count, ESR or CRP) (120). The TICORA study threshold was moderate disease activity (DAS44 > 2.4) or higher. The CAMERA study threshold was monitored by a computer programme and comprised less than 20% improvement in swollen joint count and less than 20% improvement in two out of a further 3 criteria (ESR, tender joint count, general well being VAS). In each study, DMARD escalation decisions for the comparator group were based upon the treating physicians clinical impression of global disease activity and, as previously described, were consistently associated with worse outcomes. This is hardly surprising since clinical examination alone for features of active inflammatory joint disease present is relatively insensitive (121,122). Hence, the effectiveness of initial step-up DMARD treatment regimens appears to be improved through formalised clinical assessments of global disease activity

The Behandel-Strategieen (BeST) Study has further demonstrated the value of strategies which aim for a predefined disease activity level (123). 508 patients with untreated RA (median symptom duration 23 weeks) were randomly allocated to receive one of four different approaches to DMARD therapy; 1. sequential monotherapy, 2. step-up combination therapy, 3. combination DMARD therapy with tapering steroid and 4. combination DMARD and tumour necrosis factor-alpha (TNF $\alpha$ ) antagonist therapy. In each treatment arm therapy was escalated until low disease activity (DAS44 < 2.4) was achieved. Patients who received the most aggressive initial treatment regimens (groups 3 and 4) experienced the fastest initial improvements in measures of disease activity and functional impairment and longer periods of sustained remission. However, at 12 months follow-up and after a greater number of treatment changes, patients in groups 1 and 2 (i.e. less aggressive initial treatment strategies), had attained similar disease activity levels and overall response rates to those in groups 3 and 4. Radiographic progression rates were significantly lesser for groups 3 and 4 and it is possible that this could partly be explained by the earlier attainment of low disease activity in the more aggressively treated groups. That is, even though the final disease activity levels were similar, the groups that experienced the fastest

improvement in inflammatory disease activity (and the lesser cumulative inflammatory burden) also demonstrated lesser rates of radiographic damage progression

To date, all strategic DMARD trials have employed low disease activity as a treatment escalation threshold rather than clinical remission. Therefore, current RA treatment guidelines state low disease activity (DAS28 < 3.2; DAS44 < 2.4) as the preferred disease activity target (124,125). However, attaining even lower levels of disease activity should be associated with better outcomes since the total overall inflammatory burden will have been lesser:

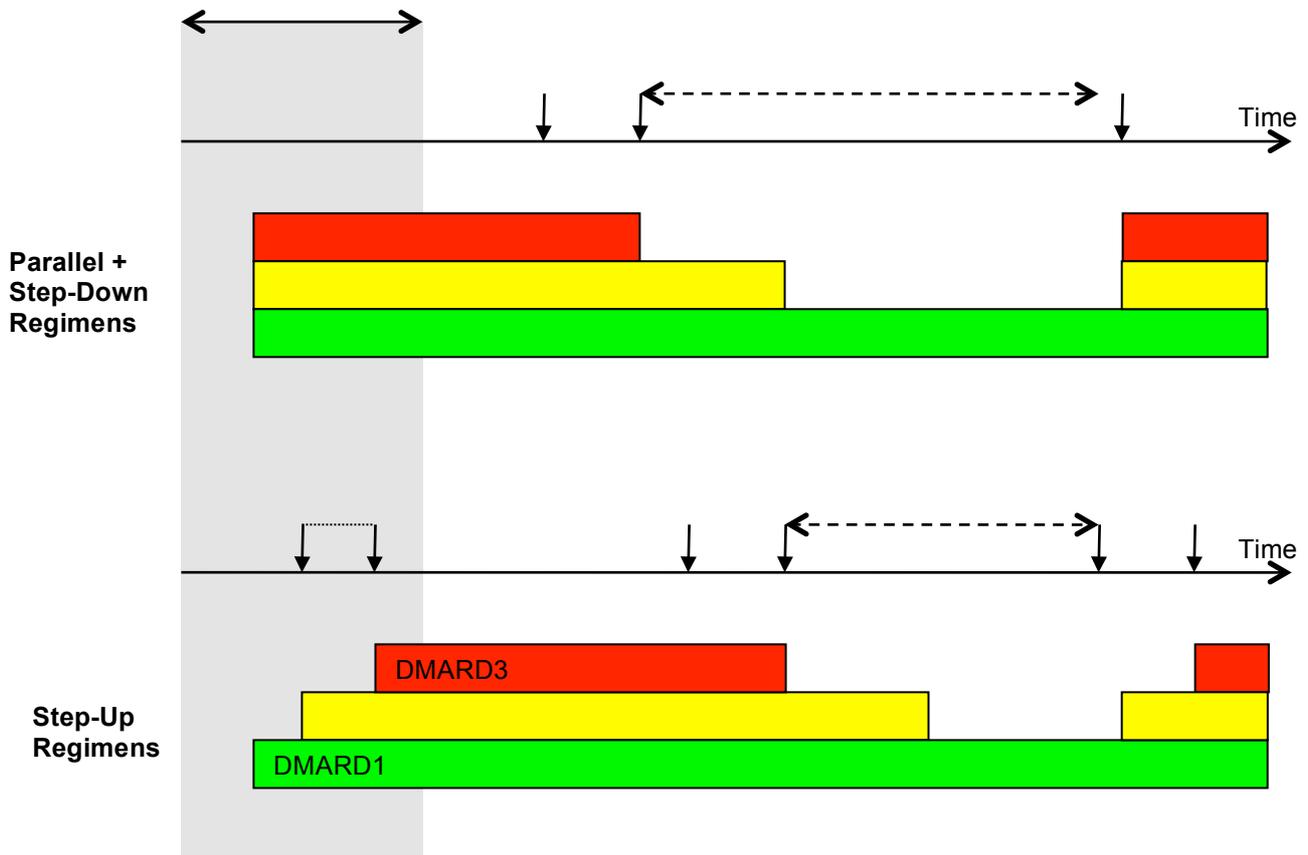
- Cohen et al compared 3 and 5 year radiological and functional ability outcomes between 30 patients in persistent remission (3 and 5 year DAS44 < 1.6) to 104 who had not achieved persistent remission (Mean 5 year DAS44 = 2.49 (i.e moderate disease activity)) (126). Compared to non-remitting patients, those who achieved sustained remission experienced greater improvements in functional ability (mean  $\Delta$ HAQ: -0.97 vs -0.65) and lesser rates of radiographic progression (mean  $\Delta$  total Sharp Score: 4.37 vs 15.01).
- A post-hoc analysis of The Active-Controlled Study of Patients Receiving Infliximab for the Treatment of Rheumatoid Arthritis of Early Onset trial (the ASPIRE trial) compared 14 and 54 week radiographic outcomes in patient's who received either methotrexate and placebo or methotrexate and infliximab who were stratified according to their disease activity level (127). At all disease activity levels, patients who received methotrexate and infliximab combination therapy exhibited significantly less rates of radiographic progression than those who received methotrexate monotherapy; though anti-TNF $\alpha$  blockers such as infliximab are one of the few available therapies that have been shown to positively influence the development of radiographic damage. After 54 weeks follow-up, and regardless of the treatment group, the level of radiographic progression was positively associated with the measured disease activity level, since there was a step-wise increase in the amount of radiographic damage accumulated from the lowest to highest disease activity groups (mean  $\Delta$  total Sharp Score: methotrexate group – remission 1.1, low SDAI 2.2, moderate SDAI 3.9, high SDAI 5.8; methotrexate and infliximab group – remission - 0.2, low SDAI -0.4, moderate SDAI 0.6, high SDAI 2.1; p values for trend not quoted)

To date, strategic studies have chosen various definitions of either low disease activity (118,123,128) or remission (119,120) as the target for DMARD therapy. Increasingly, consensus statements and international guidelines are advocating that DMARD therapy should be steered to achieve either clinical remission (111,124) and/or imaging remission (129), the presumption being that complete abrogation of inflammatory activity will be associated with the lowest likelihood of progressive disease. However, it is important to consider that there remain several different methods of classifying remission, that are not interchangeable and do not recognise identical disease states (130). Further, the risk:benefit balance of achieving modern definitions of remission remains to be determined and it may be that pursuing increasingly lower levels of disease activity is associated with either minimal additional clinical benefits and/or a higher risk of treatment associated adverse effects.

### 1.2.2 DMARD Treatment Regimens

The preceding section describes general principles that underpin an aggressive management approach that aims to optimise early treatment responses in newly diagnosed RA. However, they do not specify either which specific DMARDs to prescribe or the order in which they can be used. Diagrammatic representations of different DMARD treatment strategies are shown in Figure 2. 'Step-up' strategies initially commence DMARD monotherapy and additional agents are added if disease activity levels remain above a treatment escalation threshold. Conversely, in 'parallel and step-down' strategies several DMARDs are commenced simultaneously and once disease activity levels are persistently below a predefined threshold, doses are gradually reduced until the disease activity level appears stable on the lowest intensity combination of DMARDs possible. Several clinical trials have attempted to demonstrate the efficacy of both these approaches:

*Step-Up DMARD Combination Therapy* – formed the basis of the DMARD strategies underpinning the TICORA and CAMERA studies. Clinical studies have demonstrated that adding hydroxychloroquine, sulfasalazine, hydroxychloroquine and sulfasalazine (131), ciclosporin (132,133), leflunomide (134) or parenteral gold (135) to methotrexate produces additional improvements in disease activity measures. Initial DMARD monotherapy remains a popular first choice amongst rheumatologists (136) since a significant proportion of patients will respond adequately to monotherapy alone (137,138) and it avoids the theoretical risks of additional adverse effects when several DMARDs are commenced simultaneously. Further, patients who respond promptly to DMARD monotherapy can experience a sustained clinical response. Analysis of the Swedish Pharmacotherapy (SWEFOT) trial demonstrated that the majority of patients who experienced a satisfactory response ( $DAS28 < 3.2$ ) to initial methotrexate monotherapy continued to exhibit sustained low disease activity over the subsequent 1 (73%) and 2 year (69%) follow-up periods (139). However, despite the group level clinical response some methotrexate responders did still exhibit deterioration in all measures of radiographic damage. Furthermore, during the initial monotherapy stage patients who will ultimately require combination therapy will remain exposed to a period of persistent disease activity which is contrary to the principle of achieving tight control as early as possible.



**Figure 2** – Diagrammatic representation of different approaches to DMARD initiation regimens in early rheumatoid arthritis  
 (Reproduced from Dale and Porter, *Best Pract Res Clin Rheum* 2010; 443-455)  
 PDAS = persistent disease activity state (DAS28 > 3.2); LDAS = low disease activity state (DAS28 < 3.2)

*Parallel and Step-Down Combination Therapy* – Theoretically, commencing two or more DMARDs with differing modes of action simultaneously could generate improved outcomes if the agents interact synergistically. However, clinical studies of parallel DMARD regimens have often had conflicting results which could partly be explained by the choice of agents and variable use of corticosteroids.

- The previously described FINRaCO study demonstrated that parallel therapy (methotrexate, sulfasalazine, hydroxychloroquine and low dose prednisolone) enabled significantly more patients to attain ACR remission (37% vs 18%,  $p=0.003$ ) and ACR50 response targets than those treated with sequential monotherapy (initially sulfasalazine) (120). Both groups experienced radiographic progression though the accumulation of new erosive damage (median change in eroded joint count: 2 vs 3,  $p=0.006$ ) and overall rate of change (median change in Larsen score 2 vs 10,  $p=0.002$ ) was significantly higher in the monotherapy group.
- Hetland et al compared the efficacy and safety of a parallel treatment strategy comprising methotrexate and cyclosporine (140). The Cyclosporine, Methotrexate, Steroid in RA (CIMESTRA) trial randomised 160 patients with untreated early RA (median disease duration 3.2 – 3.9 months) to receive either methotrexate and cyclosporine combination therapy or methotrexate monotherapy. Over the course of the study, cumulative intra-

articular steroid doses were similar between both groups. After 52 weeks follow-up, response rates favoured the combination therapy group with a significantly greater number demonstrating an ACR20 response (85% vs 65%;  $p=0.02$ ). The proportion of patients achieving ACR50 and ACR70 responses was higher in the combination therapy group but did not achieve statistical significance. There was no significant radiographic progression observed within or between either treatment group. Furthermore, a greater proportion of the combination therapy group experienced either hypertrichosis (33% vs 8%,  $p<0.001$ ) or a greater than 30% increase in serum creatinine (19% vs 6%,  $p=0.03$ ).

In contrast, two earlier randomised trials demonstrated no, or modest, benefit of methotrexate and sulfasalazine combination therapy over either agent as monotherapy:

- Haagsma et al compared 12 month clinical outcomes between 105 untreated, early RA patients (<12 month's duration) randomised to receive either methotrexate monotherapy, sulfasalazine monotherapy or methotrexate and sulfasalazine combination therapy (141). Overall, patients who received combination therapy tended to experience slightly higher improvements in measures of clinical disease activity and functional ability. However, there were no statistically significant between group differences and patients in the combination therapy group did experience a higher rate of gastrointestinal intolerance
- Dougados et al compared 12 month clinical and radiological outcomes between 205 untreated, early RA patients (<1 year's duration) randomised to receive either methotrexate monotherapy, sulfasalazine monotherapy or methotrexate and sulfasalazine combination therapy (142). Patients who received combination therapy experienced significantly greater improvements in DAS44 scores (mean change: SASP -1.15, MTX -0.87, MTX+SASP -1.26,  $p=0.019$ ) and tender joint counts (mean change in RAI: SASP -7.1, MTX -4.1, MTX+SASP -9.4,  $p=0.001$ ) but not any other disease activity measure. Follow-up x-rays demonstrated a similar rate of radiographic progression between all groups. Patients within the combination therapy group experienced significantly higher rates of gastrointestinal intolerance and liver function test abnormalities.

Having commenced parallel combination therapy it is logical to consider eventually either reducing the dose, or withdrawing at least some of the component DMARDs to minimise the number of medications a patient requires. The Combinatietherapie Bij Reumatoide Artritis (COBRA) trial was the first to systematically investigate the efficacy of this approach (143). One hundred and fifty-six patients with early, active RA (median duration 4 months) were randomised to receive either methotrexate, sulfasalazine and high-dose tapering prednisolone parallel therapy or sulfasalazine monotherapy. In the parallel therapy group prednisolone, and then methotrexate, doses were tapered to cessation at set time points rather than whenever disease activity thresholds were attained. Within the first 28 week follow-up period, and particularly whilst they remained on prednisolone, patients within the parallel therapy group experienced significantly greater improvements in all measures of disease activity. However, the discontinuation of

prednisolone coincided with a gradual coming together of both groups' outcomes such that there were no significant clinical difference between either group after 56 weeks follow-up. The initial rapid improvement in disease activity measures experienced by the parallel therapy group was associated with initial slowing in radiographic damage progression. Over the first 28 week follow-up period (during parallel therapy) the parallel therapy group experienced a significantly slower rate of radiographic damage progression (median  $\Delta$  total Sharp Score: 1 vs 4,  $p < 0.0001$ ) which persisted, though the difference was less marked, over the second 28 week follow-up period (median  $\Delta$  total Sharp Score: 1 vs 2.5;  $p = 0.04$ ).

Both the FINRaCO and COBRA trials demonstrated an apparent benefit of parallel therapy over sulfasalazine monotherapy, and both incorporated prednisolone into the initial DMARD combination. In the COBRA study the initial apparent clinical benefits of parallel therapy was lost after the discontinuation of prednisolone. Furthermore, in the trials reported by Haagsma et al and Hetland et al parallel therapy with methotrexate and sulfasalazine together was only marginally better than monotherapy with either agent. Therefore, it is possible that the initial, benefits of parallel therapy demonstrated in the FINRaCO and COBRA trials are in fact related to the rapid immunomodulatory effects of corticosteroids rather than the theoretical synergistic effects of using simultaneous DMARDs

Compared to DMARD monotherapy, any treatment strategy that includes DMARD combination therapy (whether as initial parallel therapy or as a step-up option) could feasibly be associated with an increased risk of adverse effects from the overlapping actions of the component DMARDs. A meta-analysis which used the results from 36 randomised DMARD trials of early and established RA has in fact demonstrated a slightly different outcome (144). Compared to DMARD monotherapy, combination DMARD therapy was shown to be significantly more effective (RR 0.35; 95% CI 0.28, 0.45). Furthermore, whilst pooled results for all DMARD combinations demonstrated a slightly higher risk of adverse effects (RR 1.37; 95% CI 1.16-1.62), combining methotrexate with either sulfasalazine and/or antimalarials (the most commonly prescribed DMARD combination) was not associated with an increased risk.

### ***Relative Efficacy of Different DMARD Introduction Regimens***

Rather than focus on specific drug combinations several studies have performed head-to-head comparisons of different DMARD introduction regimens in an attempt to demonstrate the relative efficacies of each approach.

- The BeST Trial randomised 508 patients with active, early RA (median symptom duration 23 – 26 weeks) to one of four different DMARD introduction treatment strategies and compared functional ability and radiographic outcomes after 1 years follow-up (123,145). The treatment groups were: group 1 sequential monotherapy (starting with methotrexate), group 2 step-up combination therapy (starting with methotrexate), group 3 combination therapy (methotrexate and sulfasalazine) with tapering high dose prednisolone, group 4 methotrexate and infliximab combination therapy. In each group, the DMARD regimen would be changed, and in most cases intensified, if DAS44 remained greater than 2.4

(moderate disease activity). Patients who received the most intensive initial therapy (i.e. groups 3 and 4) required the fewest therapy adjustments, experienced earlier and faster improvements in functional ability and clinical disease activity measures and were more likely to have achieved sustained low disease activity. Furthermore, high proportions of patients in group 3 (78%) and group 4 (50%) had been able to discontinue their most potent immunomodulatory therapy (group 3 prednisolone, group 4 infliximab) because of sustained low disease activity. After 1 and 2 years follow-up, there were no significant between-group differences in HAQ scores, ACR20 and ACR70 response rates and the proportion of patients achieving clinical remission, suggesting that, within a treatment strategy aiming for low disease activity (i.e. tight control), step-up non-biologic DMARD strategies can eventually be as effective as initial combination strategies that incorporate powerful immunomodulatory agents. However, despite the apparent similarities in clinical response, initial parallel combination therapy was associated with superior radiographic outcomes. Plain radiographs of patients who received parallel combination therapy demonstrated significantly less progression of the total radiographic damage score (mean change in modified-Sharp score: 9.0 vs 5.2 vs 2.6 vs 2.5,  $p=0.005$ ), erosions score (mean change: 4.7 vs 3.1 vs 1.1 vs 1.3,  $p<0.001$ ) and joint space narrowing score (mean change: 4.3 vs 2.1 vs 1.5 vs 1.2). Furthermore, fewer patients demonstrated severe radiographic progression (change in modified-Sharp score  $> 20$ : 18 vs 7 vs 1 vs 1).

- Like the BeST trial, the Triple Therapy in Early Active Rheumatoid Arthritis trial compared a step-up DMARD introduction strategy to initial parallel combination therapy; however, the observed outcomes were noticeably different (146). Patients with newly diagnosed (mean duration 10–13 months), untreated RA were randomised to receive either step-up DMARD therapy (sulfasalazine monotherapy >> sulfasalazine and methotrexate >> sulfasalazine, methotrexate and hydroxychloroquine) or initial parallel combination therapy (all three agents). Patients were reviewed monthly (unlike BeST where patients were reviewed 3 monthly), received combinations of intra-articular and intra-muscular corticosteroid injections and DMARD regimens were steadily intensified until low disease activity ( $DAS28<3.2$ ) was achieved. After 12 months follow-up, both groups had experienced similar improvements in measures of disease activity and functional ability and, whilst there was a trend in favour of step-up therapy, there were no statistically significant differences in either the EULAR response rates or clinical remission. Radiological outcomes differed from those reported by the BeST trial. Patients in both groups experienced a similar, small amount of radiological progression (mean change in total Sharp score: step-up 6.0 vs parallel 6.6, 95%CI -3-2) with no significant between-group difference in the change in total Sharp score, erosion score (mean change: 1.1 vs 1.7, 95%CI -1.5,0.3) or joint space narrowing score (mean change: 4.9 vs 4.8, 95%CI -2,2).

Two further randomised, strategic trials have recently described whether different variations of approach to step-up therapy, in patients who had experienced an inadequate response to

methotrexate monotherapy, might have significant impacts upon patients short-medium term outcomes

- The Swedish Pharmacotherapy (SWEFOT) trial investigated whether the components of step-up DMARD combinations might influence clinical outcomes in patients who have experienced an inadequate response to initial methotrexate monotherapy (147). 487 patients with early (symptom duration < 1 year), untreated RA were initially treated with methotrexate monotherapy. 258 patients (53%), who did not achieve an adequate response to methotrexate monotherapy (DAS28 > 3.2), were then randomised to 'step-up' to combination therapy with either methotrexate, sulfasalazine and hydroxychloroquine or methotrexate and infliximab. After 12 months further follow-up, adding infliximab allowed significantly more patients to achieve an EULAR good response than adding sulfasalazine and hydroxychloroquine (47% vs 32%,  $p=0.0107$ ). After 24 months follow-up there was no significant between group differences for clinical measures of treatment response (EULAR Good response: 31% vs 38%,  $p=0.204$ ) though patients who received infliximab did demonstrate significantly lower rates of radiographic progression (mean  $\Delta$  total Sharp score: DMARD group 7.23 vs Infliximab group 4.0,  $p=0.009$ ). Short and medium functional ability measures have not yet been reported.
- Moreland et al investigated whether delaying commencement of combination DMARD therapy would have any negative impact on clinical and radiological outcomes in patients who had an initial inadequate response to methotrexate monotherapy (148) (149). 755 patients with early RA (mean duration 3.6 months) were randomised to receive either methotrexate monotherapy or parallel combination therapy (methotrexate, sulfasalazine and hydroxychloroquine or methotrexate and etanercept). After 24 weeks, the threshold for stepping up to combination therapy was persisting moderate to high disease activity (DAS28  $\geq 3.2$ ). Twenty eight percent of patients in the methotrexate monotherapy group achieved low disease activity and therefore did not require combination therapy. Over the 2 years follow-up period, and following commencement of combination therapy, the clinical and radiographic outcomes of patients who received delayed combination therapy were virtually indistinguishable from those who had received immediate combination therapy.

Taken together the results of the preceding four studies suggest that a pragmatic approach to the use of DMARD therapy in early RA would use a step-up introduction strategy, aiming for at least low disease activity, and based along tight control principles. The BeST and TEAR trials both demonstrate that the initial clinical response to step-up combination therapy is similar to initial parallel therapy. Patients in both parallel treatment groups of the BeST study did exhibit better radiological outcomes than those who received initial step-up therapy. This might partly be explained by the initial rapid improvements in measures of inflammatory disease activity; however, it is not possible to correct for the direct disease modifying effects of the associated prednisolone and infliximab. The available results from the trial conducted by Moreland et al suggest that delaying introduction of combination therapy does not necessarily disadvantage

those patients who fail to respond to methotrexate monotherapy. Delayed combination therapy also avoids overtreatment of the significant proportion of patients who will experience an adequate initial response to DMARD monotherapy. The timing of biologic therapy remains unclear. In the BeST study, patients who received combination DMARD and tapering prednisolone achieved similar clinical and radiological outcomes to those who initially received methotrexate and infliximab. In the SWEFOT study, patients who failed to respond to methotrexate monotherapy achieved significantly better 1 year clinical outcomes if infliximab was added instead of additional conventional DMARDs though clinical response rates after 24 months treatment were similar. However, the trial reported by Moreland et al comments that the results for both immediate combination therapy groups were similar. Further, such early use of biologic therapies may prove excessively expensive and be restricted by individual countries prescribing authorities. Importantly, neither the SWEFOT trial nor Moreland et al's trial have yet reported whether patients who experienced an initial adequate clinical response to methotrexate monotherapy also experienced similar medium-long term clinical, radiological and functional outcomes to those patients who received a more aggressive treatment regimen.

### **1.3 Global Disease Activity Assessment**

As a chronic, inflammatory condition, which has the potential to affect any of a large number of synovial joints, the clinical presentation and course of RA can be highly variable. Traditionally, decisions regarding the need to change DMARD therapy were based upon largely subjective interpretations of the level of active disease present; such as, the patient's description of recent symptoms, identification of clinically synovitic joints and relative changes in laboratory markers of the acute phase response. It is now routine for clinical trials in RA to describe changes in global RA disease activity through the use of composite scores, which integrate several commonly recorded patient reported measures, clinical examination findings and laboratory results, to generate a single numerical value. Furthermore, several of the previously described DMARD treatment strategy trials have used composite disease activity measures as either thresholds to trigger changes in DMARD therapy and/or outcome targets. Thus, increasingly, most RA treatment guidelines recommend the regular assessment of composite disease activity measures and their use has gradually filtered into routine clinical practice (111,112,125).

#### **1.3.1 Clinical examination**

Until DMARD strategy trials started to demonstrate the benefit of using composite measures of disease activity, treatment change decisions were informed by clinicians either identifying the presence of synovitis during clinical examination, or judging that the symptoms, signs and laboratory findings evident at a single time point represented 'active' RA. Unfortunately, in many instances, clinical examination alone has proven to be relatively insensitive (i.e. misses some areas of active disease), non-specific (i.e. misclassifies or misinterprets clinical findings) and could lead to erroneous treatment decisions. In the TICORA trial, DMARD escalation decisions in the routine therapy group were based on the clinician's interpretation of their examination findings (118) and it was clearly shown that this approach lead to worse clinical and radiological outcomes. It's worth noting that, whilst the intensity of follow-ups visit was lesser, the treatment escalation protocols were the same for both routine and intensive treatment groups.

Imaging studies have shown that identification of synovitis through clinical examination alone is insensitive, usually underestimates the amount of synovitis present, and, by extension, could leave patients at risk of persistently active, low-level synovitis (i.e. undertreatment in the presence of active disease):

- Wakefield et al performed simultaneous clinical and musculoskeletal ultrasound (MSUS) examination on 80 patients with untreated inflammatory oligoarthritis (mean duration = 18 weeks) (122). Compared to clinical examination, MSUS examination identified a higher overall number of joints with ultrasonographic evidence of synovitis (clinical examination evidence of synovitis = 12.6% vs MSUS evidence of synovitis = 27.5%) and demonstrated that 13% of clinically asymptomatic joints also had evidence of subclinical synovitis. MSUS evidence of subclinical synovitis, indicating wider spread joint involvement than suspected on clinical grounds, led to the majority (58%) of patients with clinically diagnosed monoarthritis being reclassified as either oligoarthritis (35%) or polyarthritis (23%)
- Szkudlarek et al compared the ability of clinical examination, MSUS and MRI to identify evidence of inflammation in the 2<sup>nd</sup> to 5<sup>th</sup> MCPj and 2<sup>nd</sup> to 5<sup>th</sup> proximal interphalangeal (PIPj) joints of 40 RA patients (150). MRI was considered the gold standard for identifying synovitis. Of the 480 joints examined, there was complete agreement between clinical and MSUS findings in 371 (77%) joints. Clinical examination identified 18 (4%) additional joints which had evidence of inflammation, whereas MSUS identified a further 91 (19%) joints with ultrasonographic evidence of inflammation.
- Filer et al performed systematic clinical and MSUS examinations of 58 patients presenting with untreated inflammatory arthritis and at least one clinically synovitic joint (151). Once again, it was shown that clinical examination alone underestimated the number of joints involved since, in every joint region examined (PIPj, MCPj, wrist, elbow, shoulder, knee, ankle), MSUS consistently identified a significantly greater proportion of joints with ultrasonographic evidence of inflammation.

Furthermore, clinical examination is a largely subjective skill that may be influenced by systematic differences in technique between individual examiners. Thus, differences of technique and interpretation between examiners may also introduce a further level of variability and inaccuracy:

- Salaffi et al compared the inter-examiner agreement of two rheumatologists who independently conducted clinical examinations of 44 early RA patients (disease duration < 2 years) (152). Examination findings demonstrated a variable (predominantly fair to moderate) level of agreement for identification of tender ( $\kappa = 0.31 - 0.62$ ) and swollen ( $\kappa = 0.20 - 0.65$ ) joints depending on the joint area being examined. Once again, systematic MSUS assessment performed on the same patients identified significantly more joints with ultrasonographic evidence of inflammation compared to those that were clinically swollen (mean number = 19.1 vs 12.6,  $p = 0.01$ )

- Stone et al determined the inter-examiner agreement of five rheumatologists conducting clinical examinations on 5 patients with RA and 5 with psoriatic arthritis (153). Different factors related to clinical examination demonstrated different levels of inter-examiner agreement. Visual identification of joint swelling showed moderate-substantial agreement ( $\kappa = 0.55 - 0.63$ ); however, there is likely to be less disagreement when overt inflammation is evident. Palpation of swelling (i.e. low grade inflammation) showed only slight-fair agreement ( $\kappa = 0.19-0.41$ ). Identification of joint tenderness, a relatively non-specific clinical sign, showed moderate agreement ( $\kappa = 0.41-0.58$ ).

The hallmark clinical examination findings for joint inflammation are joint tenderness and/or joint swelling and these are important components of the most commonly used composite disease activity scores. Unfortunately, these findings may also be present when RA overlaps with other conditions associated with joint pain. Thus clinicians may potentially misattribute clinical findings to RA activity and use this to erroneously justify DMARD escalation decisions. Such decisions may have little clinical/symptomatic benefit but still place the patient at risk of adverse effects (i.e. an adverse risk:benefit ratio). Wolfe and Michaud have proposed that a significant minority (17.1%) of patients develop an overlap syndrome between RA and fibromyalgia, a chronic pain condition associated with increased levels of pain, disability and fatigue (154). In two similar clinical, established RA cohorts, Pollard et al have reported a prevalence of fibromyalgic-RA between 12% to 17% (155). Fibromyalgic-RA patients were identified by disproportionately elevated tender joint counts (difference between tender and swollen joint counts  $> 7$ ) and exhibited measures of pain, fatigue and functional ability that were consistently worse than a non-fibromyalgic-RA comparator group. Furthermore, patients with fibromyalgic-RA exhibited significantly higher DAS28 scores (mean: 5.7 (95%CI 5.3-6.1) vs 4.0 (95%CI 3.7-4.3)) based largely upon higher tender joint counts (mean: 16 (95%CI 14-18) vs 4 (95%CI 3-5) and global visual analogue (mean: 61 (95%CI 53-68) vs 37 (95%CI 31-42) scores but similar swollen joint counts (mean: 3 (95%CI 2-4) vs 4 (95%CI 3-4) and ESR (mean: 33 (95%CI 22-43) vs 28 (95%CI 24-32). Hence, if the physical examination findings to calculate the DAS28, and using this as part of a tight-control DMARD treatment strategy aiming for low disease activity, had been used to justify DMARD escalation a significant minority of patients would face being treated with higher doses, or increasingly complex combinations, of potentially toxic DMARDs which were unlikely to be effective in treating their painful joint symptoms.

### **1.3.2 Acute Phase Reactants**

Theoretically, the laboratory measures of the acute phase response (e.g C-reactive protein (CRP), erythrocyte sedimentation rate (ESR)) could prove useful additional measures of RA disease activity since they remove any potential bias associated with the subjective interpretation of physical examination findings. However, acute phase markers are inherently non-specific and can be influenced by external factors such as intercurrent illness. Furthermore, acute phase markers are of limited value in the substantial subset of patients (approximately 30-40%) who have clinically active disease but fail to mount a measurable elevation in acute phase reactants (156,157). Cohort studies have demonstrated that elevation of inflammatory markers is

associated with elevation of other disease activity measures and an increased risk of radiographic progression:

- Dixon et al compared the levels of various acute phase markers (CRP, ESR, haptoglobin, fibrinogen) to clinical measures of global disease activity (articular index) in 105 RA patients treated with a variety of older non-biologic DMARDs (158). Serial levels of all markers fell significantly in response to commencement of therapy and, for each treatment group, showed moderate to strong positive correlations with corresponding clinical articular index scores (mean correlations: 0.774 – 0.954,  $p < 0.01$  -  $< 0.001$ )
- Van Leeuwen et al demonstrated that time-integrated acute phase reactant levels correlate positively with radiographic progression over the same period (68). 110 patients with early RA (mean symptom duration < 26 months) underwent monthly measurement of acute phase reactant levels and 6 monthly radiographs of hands and feet for 36 months. At all follow-up time points, cumulative CRP and ESR levels correlated positively with the observed amount of radiographic progression (Spearman's correlation coefficient: CRP 0-12 months  $r=0.599$ ; 0-24 months  $r=0.607$ ; 0-36 months  $r=0.638$ ;  $p < 0.001$ . ESR 0-12 months  $r=0.522$ ; 0-24 months  $r=0.498$ ; 0-36 months  $r=0.507$ ,  $p < 0.001$ ). Thus, persistent elevation of inflammatory markers is associated with ongoing radiographic progression

### **1.3.3 Composite Disease Activity Measures**

Individual clinical and laboratory variables have been shown to perform poorly as single markers of global disease activity since they each correlate only moderately with other disease activity markers (159). In the same analysis, combining several of the individual variables into a composite measure greatly enhanced the validity of the resultant disease activity score (mean Pearson correlation coefficient: DAS44 0.63; Mallaya Index 0.65, Riel Index 0.61). Several composite scores have been proposed (summarised in Table 1) and each lends differing weight to various combinations of patient reported outcomes, clinical findings and laboratory results to generate a numerical score that attempts to objectively represent global RA disease activity (160-167). The numerical output of each measure allows disease activity levels at a single time point to be categorised according to severity (remission, low, moderate and severe) and allows serial measurements to 'track' fluctuations in disease activity levels over time, or in response to a therapeutic intervention. As one component of a tight control treatment strategy, composite disease activity measures have been shown to contribute significantly to improved clinical outcomes in early RA (118,128). Indeed, all recent national and international early RA management guidelines have recommended the use of some form of composite disease activity measure(112,124,168-170). However, despite these recommendations, a recent survey of 335 American Rheumatologists has shown that whilst the majority of respondents felt composite measures were useful in clinical practice (48-75%), only a minority were actually using composite measures regularly (DAS44 5.4%, DAS28 27.8%, SDAI 6.6%, CDAI 15.2%, PAS 6.9%, PAS-II 1.8%, RAPID3 29.25%, RADA15 1.19%) (165)

Composite disease activity scores can be categorised based upon the types of variables they employ:

*Patient Reported Composite Outcome Measures* – patient reported outcomes are based solely upon patient's own assessments of their symptoms and functional ability. Hence, they are generally easy to use, quick to perform and provide a patient focussed measure of disease activity. However, they do not incorporate a clinical assessment of disease activity (notably the presence/absence of clinical synovitis) and therefore may be influenced by external factors not directly related to RA disease activity (e.g secondary degenerative joint disease, comorbidities). Furthermore, responsiveness to change and long-term predictive power have often not yet been established

- Patient Activity Scores (PAS / PASII) (171) – are calculated using patient reported 10cm visual analogue scores for pain and global health and either HAQ or HAQ-II (PAS-II) questionnaires. The PAS-II has been validated in a wide range of rheumatological conditions where as the PAS has only been validated in RA. PAS-II correlates fairly with DAS28 ( $\kappa = 0.29$ ) and CDAI ( $\kappa = 0.40$ ) (172); however, the longitudinal performance of either measure in response to changes in disease activity has not yet been studied
- Routine Assessment Patient Index Data (RAPID3) (173) – The RAPID3 measure is the most commonly used of the proposed RAPID measures and comprises patient reported 10cm visual analogue scores for pain and global health and the MD-HAQ questionnaire. At a single time point, RAPID3 scores correlate moderately-strongly with DAS28 (Spearman's correlation coefficient: 0.39 – 0.61) and CDAI (Spearman's correlation coefficient: 0.54 – 0.77) (174); however, its responsiveness to changes in disease activity has not yet been studied
- RA Disease Activity Index (RADAI) (7,175) – The RADAI questionnaire uses quite different components to other patient reported disease activity measures. Patients respond to 5 questions focussed on their perception of RA symptoms over the preceding 6 months. RADAI correlates positively with other composite disease activity measures at a single time point (Spearman's correlation coefficients: DAS28, SDAI, CDAI = 0.64-0.74;  $p < 0.001$ ) (176). Similarly, changes in RADAI correlate strongly with corresponding changes in DAS28 ( $R^2 = 0.70$ ,  $p < 0.0001$ ) (177).

	Components						Scoring				
	Tender Joint Count	Swollen Joint Count	Patient Global VAS	Physician Global VAS	ESR / CRP	Range	Low	Moderate	High	Remission	
<b>DAS28</b> 28 Joints Disease Activity Score	28 joints	28 joints	✓	X	✓	0 – 9.4	≤ 3.2	> 3.2 AND ≤ 5.1	> 5.1	< 2.6	
	0.56√(28TJC) + 0.28√(28SJC) + 0.7lnESR + 0.014(patient global VAS)										
<b>DAS44</b> 44 Joints Disease Activity Score	RAI*	44 joints	✓	X	✓	0 – 10	≤ 2.4	> 2.4 AND ≤ 3.7	> 3.7	< 1.6	
	0.54√(RAI) + 0.065(44SJC) + 0.33ln(ESR) + 0.0072(patient global VAS)										
<b>SDAI</b> Simplified Disease Activity Index	28 joints	28 joints	✓	✓	✓	0.1 – 86.0	≤ 11	> 11 AND ≤ 26	> 26	< 3.3	
	28TJC + 28SJC + patient global VAS + physician global VAS + CRP										
<b>CDAI</b> Clinical Disease Activity Index	28 joints	28 joints	✓	✓	X	0 – 76	≤ 10	> 10 AND ≤ 22	> 22	< 2.8	
	28TJC + 28SJC + patient global VAS + physician global VAS										
<b>PAS / PASII</b> Patient Activity Score	HAQ (PAS) HAQ-II (PASII)	✓	✓	Pain VAS	X	0 - 10	< 3.7	≥ 3.71 AND < 8.0	≥ 8.0	< 0.26	
	1/3(3.33(HAQ or HAQII) + pain VAS + patient global VAS)										
<b>RADAI</b> RA Disease Activity Index	Patient Completed Questionnaire						0 – 10				Changes greater than 1.4 RADAI points are considered significant
	Mean of 5 component scores										
<b>RAPID3</b> Routine Assessment Patient Index Data	MD HAQ	✓		Pain VAS	X	0 – 10	≤ 2.0	> 2 AND ≤ 4	> 4	< 1	
	1/3(3.33(MD-HAQ)+ pain VAS + patient global VAS)										

**Table 1** – Summary of components of RA composite disease activity measures and scoring ranges  
VAS – 10cm Visual Analogue Scale, RAI – Ritchie Articular Index, SJC – Swollen Joint Count, TJC – tender joint count

### *Patient and Clinician Composite Outcome Measures*

- Clinical Disease Activity Index (CDAI) (178) (179) – The CDAI is calculated using clinical examination findings and patient reported outcomes and, since it does not include laboratory variables, can be available immediately at the point of care. Indeed, the CDAI was derived by demonstrating that acute phase reactants only contributed a small proportion to the variance of other composite disease activity measures (DAS28 15%; SDAI 5%). CDAI scores correlate strongly with other composite disease activity measures (DAS28  $R=0.89-0.90$ ; SDAI  $R = 0.90-0.91$ ) and demonstrate a similar relationship to HAQ as other composite measures ( $R = 0.45-0.47$ ). Longitudinal assessments have demonstrated that changes in CDAI disease activity are similar to changes in DAS28 and SDAI disease activity and that changes in CDAI scores correctly categorise patients into appropriate ACR response groups. Furthermore, CDAI appears to have similar predictive properties for radiographic progression to DAS28 and SDAI since time-averaged scores for each measure showed similar correlations to changes in radiographic damage after 36 months follow-up ( $R$  coefficient between time averaged measure and change in Larsen score: DAS28 0.58 (95%CI 0.37-0.73); SDAI 0.59 (95%CI 0.39-0.74); CDAI 0.54 (95%CI 0.32-0.70)) (178)

### *Patient, Clinician and Laboratory Composite Outcome Measures*

The Disease Activity Scores and Simplified Disease Activity Index utilise similar clinical, laboratory and patient reported measures to generate numerical outputs. Both use tender and swollen joint counts, patient global health 10cm visual analogue scores and a measure of the acute phase response. The Simplified Disease Activity Index also includes a physician global health 10cm visual analogue score. The Disease Activity Scores employ complex mathematical calculations to apply differential weightings of importance to individual variables whereas the Simplified Disease Activity Index is simply the sum of the component variables and therefore is relatively simple to calculate.

- Disease Activity Scores- The 28 and 44 joint Disease Activity Scores (180) (DAS28 and DAS44 respectively) are the most commonly used composite measures of RA disease activity. They have been extensively validated both as clinical trial outcome measures and tools to measure response to therapeutic intervention (161) with the DAS28 often being considered quicker and more convenient to administer. Whilst both scores employ slightly different variables and scoring ranges, for given patients their outputs show strong correlation ( $R=0.97$ ) (180). For calculation purposes, either the CRP or ESR can be used to represent the level of acute phase response; however, the outputs are not interchangeable since DAS28-CRP scores tend to be lower than equivalent DAS28-ESR scores (181). Acute phase reactants contribute significantly to the final DAS28 value which, in some instances, can lead to erroneous categorisation of disease activity since reactants levels may fall in response to therapy but patients may still exhibit swollen, synovitic joints (182). Conversely, external factors may stimulate increased acute phase reactant levels that contribute to an apparently elevated DAS28 score that is not related to active RA. At a

single time point, and over time, persistent elevation of DAS44 scores are the largest contributor to functional decline (HAQ) (183). Additionally, persistent elevation of DAS28 over time correlates positively with radiological progression (Change Larsen Score:  $R=0.58$  (95CI 0.37-0.73),  $p<0.001$ ) (178)

- Simplified Disease Activity Index (SDAI) – The SDAI (184) is simpler to calculate than the DAS28, though comprises similar variables. Balanced weighting of individual variables generates an output which correlates positively with DAS28 ( $R=0.8-0.92$ ;  $p<0.0001$ ) (185) though a SDAI score less than 3.3 dictates a more stringent description of remission. Hence, SDAI remission is one of two definitions of remission proposed in the recent ACR/EULAR Boolean Definition of Remission for Clinical Trials (186). Following treatment, changes in SDAI reflect corresponding changes in other disease activity measures and persistent elevation of SDAI is positively associated with an increased risk of experiencing radiographic progression (184). Whilst SDAI (sensitivity 90%; specificity 86%) levels outperform DAS28-ESR (sensitivity 87%; specificity 70%) and DAS28-CRP (sensitivity 86%; specificity 78%) at predicting which patients require DMARD therapy changes (187), the effectiveness of SDAI to ‘steer’ DMARD therapy has not yet been formally assessed in a ‘treat-to-target’ treatment strategy study.

Despite each of the composite measures utilising similar components of the ACR core-set variables, different mathematical constructs can lead to each measure categorising patients differently. Overall, DAS28 and SDAI appear to categorise patient’s disease into similar disease activity level groups (130,188). In a longitudinal study of 200 early RA patients receiving DMARD therapy, DAS28 showed good agreement with both SDAI and CDAI at identifying low disease activity (i.e the treatment target) during follow-up visits ( $\kappa = 0.68$  and  $0.67$  respectively) whilst SDAI and CDAI showed excellent agreement ( $\kappa = 0.97$ ) (130). In the same study, DAS28, SDAI and CDAI demonstrated lesser agreement at identifying clinical remission ( $\kappa = 0.48$  and  $0.52$  (moderate) respectively), whereas SDAI and CDAI still demonstrated excellent agreement ( $\kappa = 0.97$ ). In fact, the SDAI provides the most stringent definition of clinical remission since patient’s swollen joint count cannot exceed 2. By contrast, patients can be categorised as meeting DAS28 remission criteria but still display up to 10 swollen joints (189,190). Indeed, comparison of different definitions of clinical remission from 2754 patients with RA demonstrated that 85% fulfilling SDAI clinical remission criteria had no swollen joints whereas only 70% of patients fulfilling DAS28 clinical remission criteria had no swollen joints (191). Taken as a surrogate for clinical synovitis, persistence of swollen joints, despite meeting clinical remission criteria, does appear to have a significant bearing on whether or not patients have actually achieved inactive RA. In a study of 114 patients treated with methotrexate monotherapy with sustained clinical remission ( $DAS28 < 2.6$  consistently for 6 months); those patients with residual joint swelling (swollen joint count  $\geq 2$ ) tended to experience greater degrees of radiographic progression than those with one or no swollen joints (mean change in Sharp/van der Heijde Score:  $2.2$  vs  $0.2$ ;  $p=0.11$ )(192). Once again, the effectiveness of SDAI at truly categorising inactive disease was demonstrated; a much smaller proportion of patients fulfilled SDAI remission criteria (46%) and, overall, this group experienced virtually no radiographic progression (mean change in Sharp/van

der Heijde Score = -0.07). Interestingly, attaining DAS28 remission and having no swollen joints appeared analogous to attaining SDAI remission since both subgroups demonstrated virtually identical rates of radiographic progression. However, whilst attainment of remission (by which ever measure) is clearly desirable, clinical scores may not fully exclude active disease. Of 93 patients who attained persistent ACR remission (a very strict definition), 13 (14%) demonstrated clinically significant erosive progression and 14 (15%) developed erosions in previously unaffected joints (193)

## **1.4 The Role of Musculoskeletal Ultrasound in the Assessment of Rheumatoid Arthritis**

Incorporation of systematic measurement of global disease activity into regular clinical assessment has undoubtedly improved short-medium term outcomes in early RA. To date, most treatment strategy trials have used an objective score, representing global disease activity, to 'steer' DMARD therapy until a predefined, acceptable lower level has been achieved. Studies have expressed their chosen target using either conventional composite scores - such as DAS44 (118) or DAS28 (128,194) – laboratory markers of the acute phase response (195,196) or a computer-performed analysis of several commonly recorded ACR core-set variables (119). However, as previously described, despite meeting existing remission criteria a subset of patients can still exhibit swollen joints (i.e clinically evident synovitis) and may be at risk of undertreatment if composite disease activity measures alone are used to guide DMARD therapy. Conversely, another subset of patients may return inappropriately elevated composite disease activity scores that are not directly related to underlying inflammatory activity but could feasibly lead to inappropriate DMARD escalation.

### **1.4.1 Persistent Disease Activity Despite Clinical Remission**

Composite disease activity scores provide a useful measure of a patient's total RA disease activity at a single time point, are a useful means of objectively assessing a patient's response to DMARD therapy changes and tracking their overall progress. However, they may be insensitive to persisting low disease activity, so leading to some patients being classified as having remission - and by extension no inflammatory disease activity - when active synovitis can be demonstrated using additional imaging modalities. Imaging studies have consistently demonstrated that despite fulfilling composite disease activity measure remission criteria a significant subset of patients still exhibit imaging evidence of active synovitis:

- Twenty-two patients with established RA underwent standardised clinical and MSUS examinations of their knees to determine differences in detection rates for common inflammatory knee lesions (197). MSUS consistently out performed clinical examination by identifying a significantly higher number of suprapatellar bursitis (39% vs 16%), knee joint effusions (61% vs 36%) and Baker's cysts (24% vs 5%). Hence, MSUS appears more sensitive than clinical examination at identifying inflammatory knee lesions.

- Eighty patients with new diagnoses of inflammatory arthritis underwent clinical and ultrasonographic assessment for the presence of synovitis (122). During targeted scanning of 459 painful, but not clinically synovitic, joints MSUS identified subclinical synovitis in 150 joints (33%). In blanket scanning of 826 asymptomatic, non-synovitic joints, MSUS still identified subclinical synovitis in 107 joints (13%). Of the 80 patients examined, MSUS identified a higher burden of joint involvement than clinical examination in 51 (64%) patients and led to 36 (29%) patients being reclassified from an oligoarthritis to a polyarthritis.
- Presence of musculoskeletal ultrasound (MSUS) and magnetic resonance imaging (MRI) evidence of inflammatory joint disease was reported in 107 established RA patients (median disease duration 7 years) whose rheumatologist classified as being 'in remission' (121). Sixty one patients (57%) fulfilled DAS28 remission criteria; of whom 48 (79%) demonstrated MSUS evidence of synovial hypertrophy; 29 (48%) demonstrated MSUS evidence of increased Power Doppler (PD) signal; 51 (84%) demonstrated MRI evidence of synovitis and 28 (46%) demonstrated MRI evidence of bone marrow oedema, a precursor of bone erosions. Furthermore, 31 (29%) patients met a very stringent definition of remission (asymptomatic patients with no tender, swollen or painful joints) but still displayed imaging evidence of inflammatory joint disease with 22 (73%) demonstrating MSUS synovial hypertrophy; 13 (43%) demonstrating increased PD signal; 25 (96%) demonstrating MRI synovitis and 13 (46%) demonstrating MRI bone marrow oedema.
- Szkudlarek et al compared the relative sensitivities, specificities and accuracy of clinical examination, MSUS and plain radiography to identify inflammatory joint lesions and damage when using MRI as a gold standard. The second to fifth metacarpophalangeal (MCP) and proximal interphalangeal (PIP) joints of 40 patients with established RA (median disease duration 5 years) (150) were examined independently using each modality. Out of 480 joints examined, clinical examination identified synovitis in 121 (25%) joints, MSUS identified synovitis in 194 (40%) joints, of which 91 (75%) weren't clinically inflamed. Furthermore, 18 (4%) joints were classified as having clinical evidence of synovitis which could not be identified by MSUS. Compared to MRI findings in corresponding joints, the sensitivity, specificity and accuracy of clinical examination at identifying synovitis were 0.40, 0.85 and 0.72 respectively. For MSUS, the sensitivity, specificity and accuracy were 0.70, 0.78 and 0.76 respectively.
- A small study of 10 patients with new diagnoses of early RA (median disease duration 10 months) described the longitudinal changes in clinical and MSUS examination findings in response to commencing infliximab therapy (198). Prior to treatment, 142 (51%) of 280 clinically examined joints appeared swollen (i.e clinically synovitic); 264 (64%) of 416 MSUS examined joints demonstrated grey scale (GS) synovial hypertrophy and 139 (33%) demonstrated positive PD signal. Following treatment, all clinical and MSUS measures improved significantly with only 7 (2.9%) clinically swollen joints remaining after 8 weeks. However, despite the clinical improvements, follow-up MSUS assessments still identified

MSUS evidence of active synovitis. After 8 weeks, 35% of clinically normal joints demonstrated GS synovial hypertrophy and 7% demonstrated positive PD signal. After 22 weeks, 178 (49%) of 364 joints examined by MSUS demonstrated GS synovial hypertrophy and 23 (7%) demonstrated positive PD signal. Nine of the treated patients achieved clinical remission; however, none achieved complete absence of MSUS abnormalities, albeit most were graded low-to-moderate activity.

- Balsa et al performed systematic clinical and ultrasonographic examinations on 97 RA patients judged to be in clinical remission on the basis of no clinical signs of active disease (199). Despite the clinical judgement, 92 (95%) patients had evidence of GS synovial hypertrophy and 41 (42%) had evidence of positive PD signal in at least one joint. Patients were then segregated depending on whether they met DAS28 and/or SDAI remission criteria (74% and 44% respectively). Fulfillment of SDAI remission criteria was more likely to identify patients with totally inactive inflammatory disease during MSUS assessment since, compared to DAS28 remission, these patients had significantly lower levels of PD signal and total PD scores.

It has been well demonstrated that, at low levels of inflammatory disease activity, clinical examination alone may be falsely reassuring and under represent the total active disease burden. Furthermore, evidence is emerging that fulfilling existing remission criteria does not necessarily equate with inactive disease. Longitudinal studies of patients fulfilling remission criteria have demonstrated disease progression thus implying that existing remission definitions do not necessarily equate to inactive disease

- Molenaar et al observed the clinical and radiographic progress of 187 established RA patients (median duration 7 years) who fulfilled ACR remission criteria at recruitment (193). After 2 years follow-up, 97 (52%) patients remained in persistent clinical remission; however, overall there was a significant increase in radiographic damage scores (median Sharp/van der Heijde score: baseline = 21, 2 years = 25,  $p < 0.001$ ). Progression rates were higher in patients who experienced an acute RA flare compared to those who remained in persistent clinical remission (median progression score: 1 vs 0,  $p < 0.001$ ). However, 7 (7%) patients in persistent clinical remission demonstrated clinically relevant progression of radiographic damage scores and 14 (15%) developed new erosions in previously unaffected joints
- A longitudinal observational study of 102 established RA patients (median duration 7 years), deemed to be in clinical remission, described how MSUS and MRI evidence of active synovitis predisposes to subsequent disease progression (200). The definition of remission was based upon clinician's judgement rather than specific criteria; however 56% of participants did meet the DAS28 definition of remission. At baseline, imaging techniques demonstrated high levels of joint inflammation. On MSUS 89% of patients demonstrated synovial hypertrophy and 63% demonstrated PD signal. On MRI, 92% demonstrated

synovitis and 53% demonstrated bone marrow oedema. After 1 year follow-up, 19% of patients demonstrated significant deterioration in radiographic joint damage scores.

- Composite disease activity scores provide an overall representation of total disease activity. However, the influence of elevation in a single component (especially swollen joint counts as a surrogate for clinical synovitis) may be attenuated if all other components are normal. Clinical outcome data for 864 RA patients (mean duration 2.4 years) treated with methotrexate monotherapy was pooled from the control arms of several recent randomised trials of biologic therapy (192). One hundred and fourteen patients (13%) were classified as exhibiting sustained DAS28 remission (DAS28 < 2.6 for 6 months), of these 77 (68%) demonstrated radiographic progression. Patients in sustained remission with one or no swollen joints (i.e. minimal clinical synovitis) demonstrated lower changes in radiographic progression scores than those with two or more swollen joints (mean change Sharp/van der Heijde Score: 0.2 vs 2.6,  $p=0.11$ ). Thus, even if clinical remission criteria are fulfilled, the persistence of clinically swollen joints is still associated with radiographic disease and may be an additional marker of active disease.

#### **1.4.2 Inappropriate Elevation of Disease Activity Measure**

The characteristic symptoms of RA joint involvement are pain, stiffness, swelling and loss of function. However, these symptoms are relatively non-specific and in RA patients their presence may not always be directly related to on-going, active inflammation. Joint pain and loss of function can be caused in RA patients by other, coexistent illnesses and therefore careful evaluation is necessary to discriminate which is the predominant cause in each individual patient's case. Coexisting painful conditions (such as fibromyalgia or osteoarthritis) can confound the outputs of composite disease activity measures, whereby elevated tender joint counts and patient reported global health visual analogue scores generate apparently elevated disease activity scores in the absence of clinical synovitis and/or an elevated acute phase response. Equally, acute phase reactants may be elevated for other reasons; for example, ESR may be elevated in the elderly or cases of dysproteinaemia (201), whereas CRP is often elevated in patients with other chronic inflammatory conditions (202). Presuming that all joint pains are related to RA could feasibly lead to inappropriate escalation of DMARD therapy, when in fact inflammatory disease is quiescent. This would be unlikely to relieve a patient's symptoms but may increase their risk of experiencing treatment side effects. Several studies have demonstrated that patients with RA might develop persistent elevation of disease activity measures despite adequate treatment of their arthritis:

- Five year follow-up of 525 newly diagnosed RA patients recruited during four different time periods between 1985 and 2005 showed that more recently recruited patients experienced a relatively milder early disease course with significantly lower mean disease activity scores for their respective follow-up periods (mean DAS28 over 5 years: 1985-1990 = 4.1; 1990-1995 = 3.1; 1995-2000 = 3.4;  $p<0.0001$ ) (203). However, despite the apparent improvements in measures of inflammatory disease activity there were no significant differences reported between pain outcomes regardless of which period patients first

commenced DMARD treatment (mean pain 10cm VAS after 5 years: 1985-1990 = 32.3; 1990-1995 = 35.5; 1995-2000 = 33.9;  $p=0.3493$ )

- Leeb et al compared the DAS28 outputs for 62 RA patients with age and sex matched fibromyalgia patients (204). Overall, there were no significant between group differences for final DAS28 scores (RA 4.23 vs fibromyalgia 4.04;  $p>0.05$ ). However, the individual variables, which contributed to the final score, did vary significantly between the diagnostic groups. RA patients exhibited significantly higher mean swollen joint counts (3.84 vs 0.04;  $p<0.0001$ ) and ESR levels (24.16 vs 8.42;  $p<0.0001$ ); whereas, fibromyalgia patients returned significantly higher mean tender joint counts (6.08 vs 12.38;  $p<0.0001$ ) and global health visual analogue scores (43.72 vs 64.21;  $p<0.0001$ ). These results emphasise the importance of determining the underlying cause of joint pain prior to making treatment decisions based on composite disease activity measures since, based upon current DMARD escalation recommendations, many of the fibromyalgia patients would have been eligible for DMARD therapy if DAS28 scores alone had been used to inform treatment decisions.
- Wolfe et al have proposed that a substantial subset of RA patients develop an overlap with fibromyalgia which contributes to them experiencing persistent joint pain and fatigue (154). 11,866 RA patients completed the Regional Pain Scale and a 10cm fatigue VAS to determine the incidence of co-existing fibromyalgia. 1731 (17.1%) respondents fulfilled criteria for fibromyalgia, were more likely to be work disabled (54.5% vs 26.4%), reported more severe symptoms across all outcome measures and had higher direct medical costs.
- Pollard et al defined 'fibromyalgic RA' as a difference of at least 7 between the tender and swollen joint counts and sought to determine its influence on the DAS28 (155). In two independent cohorts 12 – 17% of patients were identified as fibromyalgic-RA. Using DAS28, fibromyalgic RA patients returned higher overall disease activity scores (5.7 vs 4.0 and 6.1 vs 4.4) and were more likely to be classified as having active disease (OR 14.3, 95%CI 5.5-37.1). Whilst tender joint counts (16 vs 4 and 18 vs 5) and global health VAS (61 vs 37 and 63 vs 46) were significantly higher for fibromyalgic RA patients, swollen joint counts (3 vs 4 and 6 vs 5) and ESR levels (33 vs 28 and 32 vs 33) were similar compared to non-fibromyalgic RA patients. These results further support the view that disproportionate elevations in tender joint counts and global health VAS may return elevated composite disease activity scores that do not necessarily reflect the level of active inflammatory disease present
- A recent questionnaire assessment of 2795 European and American patients with established RA (mean disease duration 9.9 years) demonstrated striking discrepancies between patient's perceptions of their RA disease control and experience of pain (205). All patients, at all levels of RA disease activity, experienced pain in some form and levels of pain were positively associated with the severity of RA disease activity. Whilst the majority of patients with mild RA reported mild pain levels (Europe 62%; USA 63%) a significant

minority were still experiencing moderate pain (Europe and USA both 36%). Furthermore, patients being treated for depression were more likely to report severe pain though the data is not sufficient to elucidate the inter-relationship between depression and self reported pain levels

- In the subset of RA patients who develop an overlap with fibromyalgia, having a diagnosis of fibromyalgia is an independent predictor of subsequent DAS28 scores. In a cross-sectional study of 270 patients with established RA, 32 (13%) also fulfilled ACR classification criteria for fibromyalgia (206). Compared to RA patients, fibromyalgic-RA patients exhibited significantly higher mean DAS28 scores (5.36 vs 4.03,  $p < 0.001$ ), tender joint counts (9.5 vs 3.0,  $p < 0.001$ ), disease activity VAS (56.5 vs 32.0,  $p < 0.001$ ), pain VAS (76.0 vs 40.0,  $p < 0.001$ ) and HAQ scores (2.0 vs 1.12,  $p < 0.001$ ) but similar ESR (29 vs 25,  $p = 0.343$ ) and swollen joint counts (3.5 vs 2.0,  $p = 0.119$ ). Furthermore, fibromyalgic-RA patients were less likely to be categorised as either clinical remission (0 vs 17%), low disease activity (3% vs 15%) and were more likely to be categorised as high disease activity (59% vs 22%). A multivariate regression analysis suggested that the additional diagnosis of fibromyalgia was an independent predictor of DAS28 score associated with a mean adjusted increase of 0.885 points.
- A substantial proportion of patients who satisfy existing definitions of remission still experience significant levels of pain suggesting either: 1. there is persisting subclinical synovitis or 2: pain is being mediated by an alternative, non-inflammatory process (e.g central sensitisation). Out of 157 patients in persistent remission (DAS28-CRP  $< 2.6$ ) 11.9-12.5% still reported clinically significant pain (MDHAQ pain  $\geq 4$ ) at each timepoint (207). Furthermore, high pain scores were positively associated with high scores for several other non-specific symptoms, including patient global assessment, fatigue, poor sleep quality, self-efficacy and quality of life. Inflammation related features, such as swollen joint counts, auto-antibody status, C-reactive protein levels and radiographic damage scores, did not show an association with pain scores suggesting that additional, non-inflammatory factors may be contributing to these patient's symptom experiences
- RA patients who first present with high total pain scores are at a greater risk of experiencing persistently high levels of pain compared to those who present with low-moderate pain scores. McWilliam et al analysed data from 1189 patients from the Early Rheumatoid Arthritis Network to describe the relationship between changes in pain and other disease activity measures over one year (208). Despite patients experiencing an improvement in disease activity measures with DMARD therapy (mean DAS28 baseline = 4.8 vs 1 year = 3.8;  $p < 0.001$ ) many (58%) continued to report incomplete improvement in Bodily Pain scores (median score: baseline = 41 vs 1 year = 51;  $p < 0.001$ ). Female sex and a high baseline DAS28-P (a measure of non-inflammatory factors comprising joint tenderness and global health VAS) were identified as independent predictors of lesser improvements in pain scores

- There is now increasing evidence that RA disease processes are associated with alteration in nociceptive responses which, in the presence of chronic painful stimuli, could feasibly lead to abnormal or altered pain sensations. Leffler et al tested pain responses around an inflamed joint and a pain-free area in patients with early (<1 year duration) and established (>5 year duration) RA (209). Both groups demonstrated allodynia to pressure around the inflamed joints. The established RA group demonstrated sensory abnormalities - reduced sensation of light touch and hyperaesthesia to minor painful stimuli (e.g innocuous cold) – that were not evident in the early RA group, suggesting the development of altered peripheral sensory processing. Furthermore, there was evidence of altered central somatosensory processing as the established group also demonstrated allodynia in non-painful areas

Taken altogether these results suggest that in a substantial subset of patients pain scores and traditional composite disease activity measures may not be accurate measures of the total inflammatory burden since: 1. changes in pain symptoms don't always follow the same trend as corresponding changes in disease activity scores, 2. additional, non-inflammatory factors may contribute to patient's pain symptoms and examination findings (especially tender joint counts and global health VAS) and 3. RA patients exhibit evidence of altered somatosensory function that may contribute to on going painful symptoms. Thus relying upon the DAS28 alone to guide treatment escalation decisions may not be appropriate since apparent elevations in certain contexts – especially when the tender joint counts and global health VAS are disproportionately elevated – may not be related to on-going inflammatory disease. In these circumstances, patients would be unlikely to benefit from further intensification of DMARD therapy and could be placed at an increased risk of experiencing adverse treatment effects but little symptomatic improvement.

## **1.5 The Significance of Inflammatory Lesions Identified by Musculoskeletal Ultrasound**

To a point, clinical examination and composite disease activity scores do provide a useful, objective measure of global disease activity which allows longitudinal monitoring and assessment of response to changes in therapy. However, in certain circumstances relying upon clinical measures alone may be either insufficiently sensitive and/or specific to allow treatment decisions to be tailored to a patient's specific needs. As the preceding sections describe two important subgroups of patients have emerged:

1. patients fulfilling clinical remission criteria with persistent subclinical synovitis since they are potentially at risk of being undertreated in the presence of demonstrably active disease
2. patients with disease activity scores elevated above DMARD escalation thresholds without active inflammatory disease since they are at risk of receiving unnecessarily intensive therapy with a poor chance of symptomatic benefit but an increased risk of drug associated adverse effects

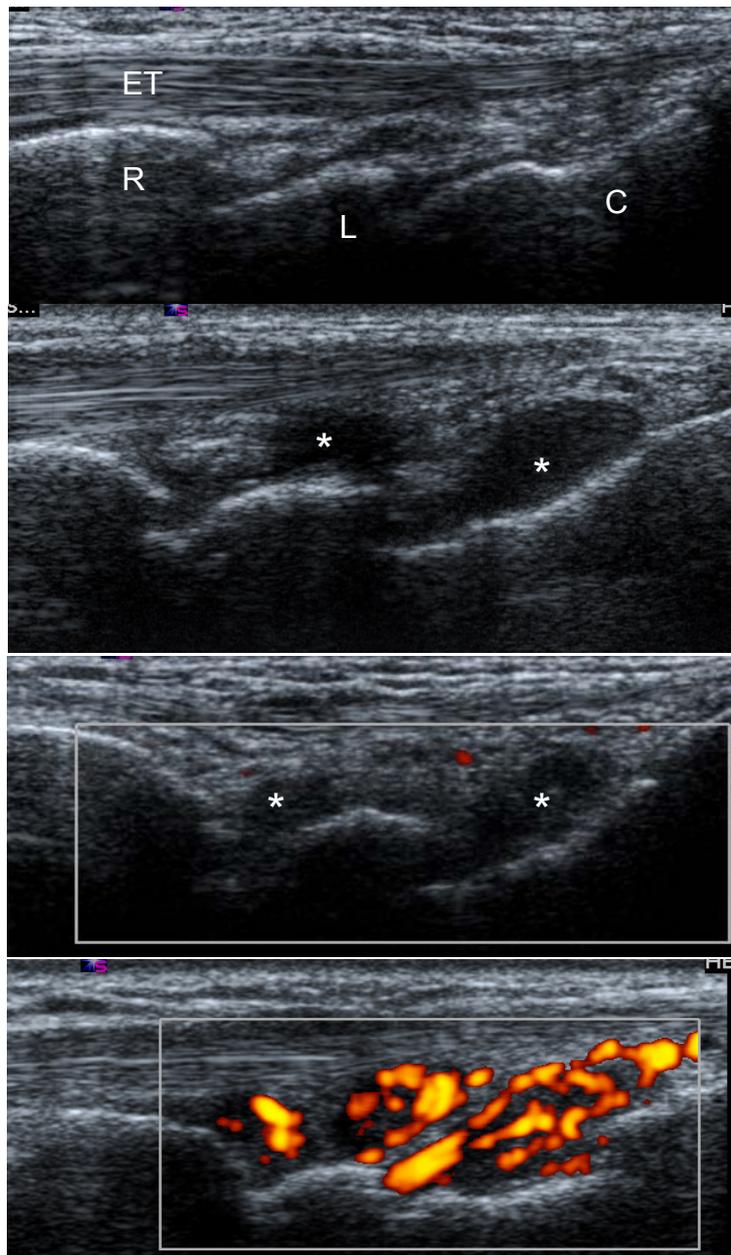
In these situations musculoskeletal ultrasound (MSUS) may offer an additional measure of global inflammatory disease that allows treatment decisions to be tailored specifically to the needs of individual patients. MSUS is being increasingly recognised as an adjunct to clinical examination since it is relatively inexpensive, reproducible, acceptable to patients (without the need for ionising radiation) and allows assessment of different joint areas during the same consultation. MSUS examination allows direct visualisation of articular and periarticular joint structures to confirm the presence or absence of inflammatory joint lesions (such as synovitis), or an alternative lesion, as the cause of a patient's symptoms. In contrast, for composite disease activity scores, several either relatively insensitive (swollen joint count) or non-specific (tender joint count, global health VAS, ESR) variables are used to generate an output that is only indirectly associated (or not) with the underlying disease process.

### **1.5.1 MSUS appearances of synovitis**

Compared to physical examination MSUS allows direct visualisation of articular and peri-articular structures and localisation of the specific lesions that may be contributing to a patient's on-going symptoms. Importantly, it specifically allows clinicians to confirm or exclude, the presence of active synovitis in patients with continuing joint pain. Histological studies have demonstrated that active synovial inflammation consists of inflammatory cell infiltration, angiogenesis / vascular proliferation and propagation of the synovial membrane (210). MSUS depicts these different characteristics using grey scale and Doppler imaging techniques. Grey scale ultrasound shows the relationship of articular structures to periarticular soft tissues and allows assessment of the size and shape of synovial hypertrophy. Colour and Power Doppler imaging allows assessment of tissue vascularity and therefore allows differentiation between active (presence of PD signals) and inactive (absence of PD signals) synovitis when applied to areas of synovial hypertrophy.

Several studies have demonstrated that MSUS representations of synovitis are accurate representations of histologically and/or MRI identified inflammatory processes.

- A small study of 10 RA patients and 10 osteoarthritis (OA) patients was one of the first to compare MSUS appearances to corresponding histological findings (211). All patients were undergoing an elective total knee replacement and underwent MSUS evaluation prior to surgery. Inflammatory pannus was identified histologically in 9 patients, 8 of whom had RA. Grey scale MSUS was less sensitive than histological assessment since both sonographers failed to identify synovial hypertrophy in all the patients with histologically evident pannus. Furthermore, grey scale imaging was less specific than histological examination with both sonographers identifying synovial hypertrophy in 4 OA patients without pannus. For both sonographers, colour Doppler sonography findings correlated much more closely to histological findings. Sonographer 1 identified colour Doppler signals in 8 of the 9 patients with histological pannus whereas sonographer 2 identified colour Doppler signals in all the patients with pannus. Furthermore, 4 patients, who displayed colour Doppler evidence of increased synovial perfusion, in the absence of pannus, did have histological evidence of active inflammation with synovial proliferation and increased vascularity



a) Normal gray scale image

R = radius  
L = lunate  
C = capitate  
ET = extensor tendon

b) Synovial hypertrophy

\* = synovial hypertrophy

c) Synovial hypertrophy without PD signal - inactive

Box = region of interest  
\* = synovial hypertrophy

d) Active synovitis

Florid PD signal (orange)  
associated with synovial hypertrophy

**Figure 3** - Examples of MSUS depictions of active and inactive synovitis  
a + b – gray scale images; c + d – gray scale images with PD

- PD assessment of knee synovium vascularity has been shown to correlate positively with histological analysis of vascularity in linked tissue specimens. Walther et al compared MSUS PD and immunohistochemical assessments of knee synovial membrane vascularity in 10 RA and 13 OA patients about to undergo total knee replacement (212). Representative PD images and tissue samples underwent additional digital image analysis to control for the subjective influence of the examiner. Strong positive correlation was demonstrated between the MSUS PD scores and the pathologists grading of vascularity (Spearman's  $\rho$  0.89,  $p < 0.01$ ). Furthermore, strong positive correlations persisted when comparing vascularity grades for digitally analysed PD images and tissue samples (Pearson's correlation coefficient  $r = 0.81$ ,  $p < 0.01$ )

- Koski et al have demonstrated that a positive synovial Doppler signal is an indicator of active synovial inflammation on histological specimens (213). MSUS and histological findings from a range of different synovial sites were compared from 44 patients with a variety of different inflammatory arthritides. Histological examination was considered the gold standard and identified abnormal synovial appearances in 43 (98%) samples, of which, 35 (79%) showed active synovitis. Combined grey-scale and PD MSUS scanning identified synovial abnormalities in 43 (98%) joints. A positive PD signal was detected in 34 of the 43 (79%) joints which showed any pathological appearances and in 29 of the 35 (83%) joints with active histological inflammation. Grades of PD signal correlated positively with levels of synovial infiltration by polymorphonuclear leucocytes ( $r = 0.397$ ,  $p < 0.01$ ) and fibrin deposition ( $r = 0.328$ ,  $p < 0.05$ ) but did not correlate significantly with either the overall histopathological score ( $r = 0.239$ ,  $p$  NS) or the vascularity grade ( $r = -0.03$ ,  $p$  NS). Thus the presence of PD signal strongly favours active synovial inflammation whereas its absence doesn't necessarily exclude it.
- MSUS and MRI show reasonable overall agreement for the identification of synovial hypertrophy and synovitis in the small joints of the hand (150). 40 RA patients underwent systematic grey scale MSUS assessment and gadolinium-enhanced MRI scanning of their dominant hand. Overall agreement between both imaging modalities was 71% with MSUS and MRI showing evidence of synovitis in 106 (38%) joints and its absence in 92 (33%) joints. Synovial hypertrophy, which was not evident on MRI, was identified by MSUS in 55 additional joints. Whereas, MRI identified synovitis in 24 joints that was not detected by MSUS. In early RA patients, MSUS was actually more sensitive than MRI at identifying synovial hypertrophy (88 vs 57 joints). Unfortunately, the paper does not report the degree of agreement between MSUS Doppler signal and post-gadolinium enhancement on MRI
- During MRI examination, administration of intravenous contrast (gadolinium), allows differentiation between areas of active and inactive inflammation. Active synovial inflammation, demonstrates post-contrast enhancement, whereas chronic synovial fibrosis does not enhance but may appear enlarged. Increased intra-articular MSUS Doppler signals appear comparable to post-contrast MRI evidence of synovial inflammation. Terslev et al compared CD signal and post-contrast MRI findings in the finger and wrist joints of 29 established RA patients (mean duration 7 years) (214). Colour Doppler MSUS and MRI agreed on the presence or absence of active inflammation in 157 joints (overall agreement 75%, kappa value 0.45). MSUS identified additional CD signals in 11 joints not evident on MRI; whereas, MRI identified enhancing synovitis in an additional 38 joints. Furthermore, CD MSUS quantitative measures of intra-articular vascularisation showed statistically significant correlations with post-contrast MRI measures of synovial thickening (MRI vs CD colour fraction  $r = 0.59$ ,  $p < 0.001$ ; MRI vs mean resistance index  $r = -0.54$ ,  $p < 0.001$ ). Thus, whilst both modalities employ different descriptors of synovial inflammation, their respective outputs have sufficient parallels to be considered at least similar. However, since the respective descriptors are related to

different underlying disease processes, the inevitable inconsistencies will prevent the outputs being fully interchangeable.

- Klauser et al have demonstrated that Colour Doppler (CD) signal is not present in the joints of healthy individuals (215). The finger joints of 46 early RA patients (disease duration < 6 months) and 10 healthy controls were graded semi-quantitatively for clinical evidence of joint inflammation and intra-articular vascularization using CD MSUS scanning. None of the joints of the healthy control subjects demonstrated detectable intra-articular CD flow signals. By contrast, 70 (25%) of joints in RA patients demonstrated some intra-articular CD flow signals. The proportion of joints demonstrating intra-articular CD flow increased as the clinical grade of inflammation increased (inactive 8%, moderately active 52%, active 58%). Administration of an intravenous bubble contrast agent significantly increased the proportion of joints in RA patients demonstrating intra-articular CD signal but had no impact on detection in healthy volunteers

### **1.5.2 Power Doppler signal corresponds to active synovitis**

Thus, an argument emerges that MSUS may provide an additional means of assessing global disease activity in RA patients based upon direct visualisation of the inflammatory lesion rather than inference from indirect clinical measures. Several overlapping strands (summarised in the preceding sections) underpin this argument:

1. at low levels of disease activity composite disease activity measures may not be sensitive enough to identify persistently active disease
2. some patients who fulfil clinical remission criteria do not have inactive disease and will exhibit evidence of disease progression
3. in a substantial subset of patients composite disease activity measures lack specificity and may remain elevated above a treatment escalation threshold even though there is no residual inflammatory disease activity
4. MSUS is more sensitive than clinical evaluation at identifying features of synovial inflammation
5. MSUS findings of synovial hypertrophy, on Grey scale scanning, and increased synovial vascularisation, on either Colour or Power Doppler scanning, compare favourably to corresponding features of synovial inflammation on histological analysis and MRI scanning

Overall, the presence of positive MSUS Doppler signals within a synovial membrane appears to differentiate well between on-going synovial inflammation and chronic fibrous synovial hypertrophy. Several studies have demonstrated that persistent Doppler signals are positively associated with observed fluctuations in clinical disease activity and predict future adverse changes in either clinical disease activity (e.g acute flare) or radiographic progression.

- The previously described longitudinal study by Brown et al compared baseline MSUS and MRI findings to clinical and radiological outcomes in 102 patients with established RA in

clinical remission (200). Clinical remission was based upon clinician's judgements; at baseline 61% fulfilled DAS28 remission criteria and 45% fulfilled ACR criteria. Over the 12 months follow-up period, patients remained in relatively stable remission with no significant changes in disease activity measures and only 5% of patients required any therapy escalation. At baseline, 89% patients had MSUS evidence of synovial hypertrophy and 63% had evidence of increased synovial PD signal. Overall, 19% of patients experienced a significant deterioration in radiographic damage scores and the risk of radiographic progression was positively associated with the presence of PD signal at baseline. Univariate regression analyses demonstrated that the total PD score in the dominant hand MCP joints was significantly, positively associated with radiographic progression in any hand or foot joint (OR 1.36, 95CI 1.02-1.81,  $p=0.036$ ). Within the MCPj, the presence of any positive PD signal (OR 12.21, 95CI 3.34-44.73,  $p<0.001$ ), the total Grey scale synovial hypertrophy score (OR 2.31, 95CI 1.06-5.52,  $p=0.032$ ) and the PD score (OR 4.4, 95CI 1.98-8.08,  $p<0.001$ ) were each significantly associated with the likelihood of developing radiographic damage. Furthermore, clinically asymptomatic joints which still exhibited positive PD signal were significantly more likely to demonstrate radiographic progression than those that had no PD signal (29% vs 4%; OR 8.77 95CI 1.54-49.89,  $p=0.014$ ).

- Peluso et al compared MSUS Grey scale and PD findings to subsequent clinical outcomes, in 48 early RA patients (mean duration 6.9 months) and 46 established RA (mean duration 118.9 months) with stable DAS44 remission ( $DAS44<1.6$  over 6 months), and demonstrated that positive synovial PD signal predicted an acute clinical flare over a 12 months follow-up period (216). Once again, high proportions of both patient groups had MSUS evidence of synovial inflammation (positive PD signal: early RA 41.7%; established RA 30.4%). 29.8% of patients experienced a clinical flare (criteria not clearly defined) over the 12 month follow-up period. A significantly higher proportion of the patients who demonstrated positive PD activity at baseline experienced an acute flare compared to those who had no PD activity (47.1% vs 20%,  $p=0.009$ ). Furthermore, patients who demonstrated an overall increase in DAS44 tended to have higher scores for synovial hypertrophy (mean 5.2 vs 2.6, no between group comparison quoted) and PD activity (mean 3.1 vs 1.1, no between group comparison quoted) compared to those who remained in clinical remission.
- Scire et al have demonstrated that positive intra-articular PD signal is the strongest predictor of subsequent disease flare in early RA patients who have already achieved DAS44 remission (217). 106 patients with early RA (mean duration 3.8 months) were treated using conventional DMARDs in a step-up tight control regimen and underwent regular clinical and MSUS evaluation for 24 months. Forty three patients achieved the studies definition of remission ( $DAS44<1.6$  on 2 occasions 3 months apart); of these 41 (95%) patients still showed MSUS evidence of synovial hypertrophy and 18 (41%) showed a positive PD signal. After achieving DAS44 remission, 14 (33%) patients experienced an acute relapse over the subsequent 6 months and exhibited significantly higher PD scores (median 1 vs 0,  $p<0.05$ ) and synovial hypertrophy scores (median 6 vs 2,  $p<0.05$ ) than

those with stable disease. Univariate logistic regression analysis demonstrated that positive PD activity in any joint was the strongest predictor of future relapse (OR 12.8, 95CI 1.6-103.5,  $p < 0.05$ ). The positive predictive value for an acute flare of PD activity was calculated as 70.6%; however, its negative predictive value of 92.3% implies that its absence is unlikely to be associated with an early acute flare

- Saleem et al used MSUS to prospectively monitor 93 RA patients who had achieved clinical remission (physician's assessment) using conventional DMARD therapy for 12 months (218). Twenty four patients experienced an acute flare and the presence of intra-articular PD signal at study recruitment was identified as the single biggest predictor of subsequent flare (OR 4.08, 95CI 1.26-13.19,  $p = 0.014$ ). Baseline MSUS findings were not quoted in the original paper.

Repeatedly demonstrating an association between MSUS evidence of subclinical joint inflammation and a subsequent increased risk of acute disease flare, and/or progression, does strongly support the presumption that the MSUS findings of active synovial inflammation (and particularly positive PD signal) do represent active disease. Further evidence to support the link between MSUS findings and disease activity can be found by demonstrating the response of MSUS representations of active synovial inflammation to changes in immunomodulatory therapy (i.e MSUS findings improve following commencement of effective therapy)

- A small study of 5 established RA patients, with clinically active disease despite DMARD therapy, demonstrated that MSUS findings of synovial inflammation improve significantly after one month's therapy with etanercept and in line with corresponding improvements of clinical and laboratory measures of disease activity (219). Patients underwent MSUS examination of all MCP joints before, and after 28 days of, treatment. Published results only relate to the right second MCPj since this showed baseline evidence of involvement in all patients. The mean number of synovial CD signals (quantitatively assessed by computer aided image interpretation) fell significantly following the administration of etanercept (mean colour signals/region of index: 23,602 to 2907,  $p < 0.001$ ) and the change in CD findings correlated well with the observed clinical change (Spearman correlation coefficient  $R = 0.85$ )
- A small observational study of 11 established RA patients (mean duration 10 years) demonstrated a significant fall in CD evidence of synovial vascularization shortly after initiating etanercept (220). MSUS was performed on the clinically worst affected joint identified during the baseline assessment. After 2 weeks treatment, all clinical and laboratory measures of disease activity had improved significantly ( $p < 0.01$  –  $p < 0.05$ ) and corresponded to significant improvements in two separate CD measures of vascularisation in the target joints (median number colour pixels per region of interest: 0.10 to 0.04,  $p < 0.01$ ; mean resistance index: 0.82 to 1.06,  $p < 0.01$ ).

- Terslev et al have also described the changes in MSUS findings of synovial inflammation immediately before, and one month after, 51 patients with established RA (mean duration 12 years) underwent intra-articular corticosteroid injections into a single, clinically inflamed joint (221). Overall, local clinical and global clinical and laboratory measures of joint swelling and disease activity improved significantly following intra-articular injection. The majority of patients also demonstrated improvement in MSUS measures of synovial inflammation. Synovial membrane volume (assessed quantitatively using pixel counting software) reduced in 38 (75%) patients (mean total pixel count: 14721 to 10169,  $p < 0.01$ ). 41 (80%) patients exhibited a significant fall in CD signal volume (assessed quantitatively using colour pixel counting software) (mean colour pixel fraction: 0.21 to 0.10,  $P < 0.001$ ) though only 32 (63%) patients experienced a corresponding increase in resistance index (mean resistance index: 0.71 to 0.79,  $p < 0.01$ ).
- Thirteen patients with established RA who had experienced an acute disease flare underwent MSUS examination of the most symptomatic either index or middle MCPj immediately before, and shortly after (within 72 hours) receiving a bolus of intravenous methylprednisolone (222). Before treatment all patients demonstrated synovial hypertrophy and increased synovial vascularisation. The majority demonstrated a significant reduction in PD signal shortly after treatment (mean percentage change in PD quantity 71%) which mirrored corresponding changes of clinical and laboratory measures and correlated significantly with the observed improved in HAQ ( $p = 0.012$ , correlation coefficient not quoted)
- Taylor et al have demonstrated that patients with RA who respond to treatment with infliximab demonstrate improvements in MSUS grey scale and PD findings which are not evident in patients who receive placebo (223). Twenty four patients with early RA (duration less than 3 years) with persisting clinical synovitis despite methotrexate therapy were randomised to receive either intravenous infliximab or placebo. Clinical and ultrasonographic (MCP 1-5 bilaterally) assessments were compared at baseline and after 18 months therapy. Baseline disease characteristics and disease activity measures were comparable between groups. Greater improvements in clinical disease outcome measures were evident in the infliximab-treated group (median change DAS28: 1.21 vs 0.39,  $p = 0.157$ ). Patients who received infliximab demonstrated significantly greater improvements in grey scale measures of synovial thickness (percentage change: 54.5 vs 13.7,  $p = 0.014$ ) and Colour Doppler signal (percentage change: 78.6 vs 26.6,  $p = 0.017$ ). Additionally, the volume of MCP synovial hypertrophy ( $r = 0.69$ ,  $p = 0.02$ ) and PD signal ( $r = 0.78$ ,  $p = 0.005$ ) present at baseline correlated strongly with progression of radiographic damage after 54 weeks follow-up in the placebo group. By contrast, patients in the infliximab group demonstrated non-significant, weakly negative correlations between baseline MCP synovial hypertrophy ( $r = -0.23$ ,  $p = 0.479$ ) and PD signal ( $r = -0.28$ ,  $p = 0.372$ ) and radiographic progression suggesting that the reduction in the volume of active MSUS evident synovitis represented a true reduction in overall disease activity as evidenced by virtually no new joint damage.

Taken together these results suggest that MSUS could have an important role to play in refining DMARD escalation decisions in carefully selected patients. MSUS allows the identification of ongoing subclinical synovitis to support decisions to escalate DMARD therapy in patients with either few clinically swollen joints, or in those who meet clinical remission criteria. Likewise, the exclusion of active synovitis in symptomatic patients should reduce patient's risk of treatment failure and/or adverse effects by discouraging unnecessary DMARD escalations. However, treatment decisions taken at a similar time point are unlikely to influence patient's long-term outcomes. It is possible that regularly assessing patient's global disease activity using MSUS could allow patients to achieve significantly better medium-long term disease outcomes through prompt identification and treatment of persistent or recurrent subclinical synovitis (even before the onset of clinical flare) and prevention of unnecessary treatment escalations when inflammatory joint disease is quiescent. Compared to patient's assessed using clinical composite disease activity measures, patients who undergo regular disease activity assessment by MSUS as part of a tight control, step-up DMARD escalation strategy, could have their DMARD regimens more closely tailored to their individual needs. Hypothetically, they should display significantly better clinical, functional and radiological outcomes (through early, aggressive suppression of persistent subclinical synovitis) and fewer DMARD associated adverse effects (through prevention of inappropriate DMARD escalation).

### **1.5.3 Implication of MSUS Disease Activity Assessment**

There is an increasing momentum within published evidence that MSUS assessment of disease activity may improve rheumatologists' ability to treat inflammatory joint disease in RA. Indeed, recent international consensus statements advocate the use of MSUS disease activity monitoring within a tight control, treat-to-target DMARD escalation regimen (129). However, before MSUS disease activity assessment is wholesale incorporated into routine practice a number of important issues should be considered:

1. So far, the bulk of published evidence supporting the use of MSUS has been observational rather than interventional. It is highly likely that, compared to DAS28, the routine use of MSUS to assess disease activity will identify a higher instance of active disease that leads to patients receiving more intensive DMARD therapy. However, so far, there have been no interventional studies that demonstrate aggressively treating subclinical disease provides improved clinical outcomes, without significantly increasing the risk of adverse events
2. MSUS is more sensitive than clinical examination for the identifying of synovial hypertrophy and active synovitis. However, as MSUS machines become more advanced and more sensitive, they may also become less specific. Earlier studies (e.g. Klauser et al) had shown that healthy subject's joints do not exhibit intra-articular PD signal (215). By contrast, several recent studies have shown that it is possible to identify physiological intra-articular Doppler signals within the joints of healthy subjects (214,224). Physiological Doppler signals may also be present in RA patients and, in the context of

RA, are likely to be interpreted as disease related rather than physiological, thus leading to the disease activity state being misclassified as active rather than quiescent and potentially leading to further DMARD escalation.

3. Unlike clinical examination, MSUS disease activity assessment requires additional equipment and time to perform. Regardless of the joints examined, performing MSUS in addition to clinical examination will require longer clinic appointments and therefore may limit the rheumatologist's ability to see as many patients within a given session. Furthermore, in addition to the initial cost of purchasing highly sophisticated equipment, the increased identification and treatment of clinical synovitis could potentially be associated with higher treatment costs since it is possible that the use of combination DMARD therapy and/or biologic therapy will be higher amongst patients assessed using MSUS that those assessed using DAS28
4. MSUS has the potential to become an extension of clinical examination and therefore, like all examination techniques, must be taught and practised. Since the interpretation of MSUS images is highly subjective there is potential for variability between individual sonographers. In fact, several validation exercises have suggested that, in experienced operators, there is moderate-good agreement between sonographers, particularly when standardised assessment methods are followed and examination is confined to easily accessible peripheral joints (152,225,226)

## **1.6 Assessment of Rheumatoid Arthritis Disease Activity and Prognosis Using Multilevel Biomarker Profiles**

Patients who develop a new inflammatory arthritis can receive diagnoses of RA through either a 'typical' clinical presentation or through fulfilment of classification criteria. However, despite the similarities in clinical phenotypes, several clinical observations suggest that the diagnostic label of RA is an umbrella term representing a heterogenous group of underlying disease processes:

1. Characteristic autoantibodies (e.g rheumatoid factors, anti-CCP antibodies) are only detectable in approximately 70% of new RA diagnoses
2. Joint involvement is not uniform at presentation. For example, whilst the majority of patients present with a symmetrical inflammatory polyarthropathy, a subset can present with an asymmetrical oligoarthritis and still fulfil RA classification criteria
3. Response to immunomodulatory therapy is heterogenous. For example, biologic therapies, targeted at core inflammatory mediators, have an average response rate of approximately 60-70%
4. Disease course after commencing DMARD therapy is highly variable even between patients with similar presenting disease characteristics and demographics

The abnormal pathogenetic and inflammatory processes which culminate in the clinical expression of an inflammatory polyarthropathy, and ultimately lead to the fulfilment of RA classification criteria, are becoming increasingly better described. Specific abnormalities relating to the development and perpetuation of RA have been described at genetic (DNA and RNA), molecular (cytokine, protein and lipid) and cellular levels both within the synovial and systemic environments (69,227). Feasibly, the heterogeneity of clinical presentation and response to DMARD therapy will be reflected by specific differences in the expression of markers relating to the activity of the underlying pathogenetic and inflammatory pathways. Ultimately, careful characterisation of an individual RA patient's disease signature (i.e how they express specific cellular and molecular markers) could allow further sub-categorisation of their illness and provide an additional means of determining their likely risk of progressive or persistently active disease and their likelihood of responding to particular immunomodulatory agents; thus, allowing highly individualised tailoring of their therapy. High through-put microarray technologies allow rapid and comprehensive characterisation of profiles at genetic, genomic (e.g RNA expression), protein and metabolic levels in a variety of different target tissues. A number of exploratory studies have described how well particular profiles correlate to specific clinical phenotypes. Broadly, profiles may provide useful additional mechanistic information relating to underlying disease processes, or may provide prognostic information estimating a patient's likely clinical course, however, their specific role in routine clinical care remains to be clarified

### **1.6.1 Expression Analysis**

The transcription of messenger RNA (mRNA) from host DNA is dynamic and can be influenced by environmental, systemic and disease related factors. At any given time point the pattern of an individual tissue's mRNA expression will provide a snapshot of which metabolic, immunological and cellular pathways are being either promoted or repressed. Microarray expression analysis experiments expose fixed genetic probes to mRNA isolated from target cells or tissue. Binding between genetic probes and tissue mRNA confirms the presence of specific mRNA segments and produces a hybridisation pattern which, when compared to a reference pattern, can be used to infer the relative up or down-regulation of specific genes. Comparing mRNA expression patterns between individuals with similar clinical phenotypes will demonstrate the degree of heterogeneity in their respective immunological and aetiological pathways. Conversely, comparing mRNA expression profiles at a group level should highlight those common patterns, or pathways, which are especially associated with a particular clinical phenotype and therefore might have either diagnostic and/or prognostic properties. To date, expression analysis experiments in RA have attempted to either 1. link mRNA profiles to specific pathgenetic processes or 2. identify mRNA profiles associated with specific clinical phenotypes

Common autoimmune diseases (RA, SLE, MS, type 1 diabetes mellitus) demonstrate similarities in gene expression profiles which allow them to be differentiated from normal controls (228). In RA specifically, several studies have attempted to link gene expression data, using a variety of different target tissues, to corresponding clinical data to either gain insight into possible pathogenetic mechanisms, describe the relationship between gene expression profile and clinical

phenotype and/or identify additional markers associated with favourable/unfavourable treatment responses.

- Van der Pouw et al performed gene expression analyses using synovial tissue collected from 15 RA patients during joint replacement surgery (229). Unsupervised hierarchical clustering demonstrated that the patients could be subdivided into two distinct groups based upon their gene expression profiles. In the first group, ten patients (Group RA-I) demonstrated increased expression of 121 genes broadly related to inflammation. This group could be further subdivided based on which specific aspect of inflammation-related gene expression was up-regulated: group RA-Ia comprised four patients with high expression of genes relating to adaptive immunity; group RA-Ib comprised 6 patients with increased expression of genes relating to classical complement pathway activation. In the second group, five patients exhibited increased expression of 39 genes relating to fibroblast differentiation but relatively low expression of genes relating to inflammation and complement activation. A crude comparison to clinical data was reported and it's notable that each subgroup contained a small number of patients. All patients fulfilled 1987 ACR RA classification criteria which led the authors to conclude that the distinct differences in gene expression profiles groupings reflected the degree of heterogeneity in RA pathogenetic processes. Statistical between-group comparisons were not reported; however, compared to other groups, Group RA-Ia exhibited higher mean ESR (38 vs 27 vs 25) and prevalence of erosive disease (100% vs 83% vs 80%) whilst Group RA-Ib had a slightly lower rate of rheumatoid factor positivity (100% vs 83% vs 100%).
- Olsen et al compared gene expression profiles in peripheral blood mononuclear cells (PBMC) between 11 early RA patients (mean duration 1.1 years), 8 established RA patients (mean duration 10.5 years) and 11 control subjects with asthma or allergic disease (230). Between group comparisons identified a gene expression pattern which was present in early RA but neither in established RA nor the control group. There was a degree of overlap between early RA expression profiles and those of patients with other autoimmune disorders (most notably SLE) suggesting that these two clinically distinct diseases might share a common pathogenetic pathway. In this study expression profiles were not compared to corresponding clinical data.
- Devauchelle et al investigated whether synovial tissue gene expression profiles could differentiate between 5 patients with established RA (mean duration 14 years) and 10 patients with osteoarthritis (OA) (231). Overall, 63 genes exhibited significant differential expression between RA and OA patients. In RA patients, 15 of the genes had higher expression levels and 48 had lower expression levels. Thirty six percent of the identified genes had known functions being related to cell cycle, signal transduction, metabolism or protease activity. Unsupervised clustering analyses correctly classified all RA and OA samples separately. Furthermore, a small validation analysis using the 63 selected genes correctly identified 2 further RA and 3 further OA patients

- Batliwalla et al demonstrated that patients with active RA exhibit different PBMC gene expression patterns to controls (232). Comparisons were performed between 29 patients with established RA (mean duration 12 years) and 21 control subjects. Hierarchical clustering analyses identified 81 genes with significantly different expression values between each group. This clustering incorporated all RA patients and three control subjects (false positives). Furthermore, a significant proportion of the up-regulated genes were related to monocyte function. How gene expression profiles were distributed amongst different RA phenotypic groupings was not described.
- Since B-cells are strongly implicated in the pathogenesis of RA, Szodoray et al compared gene expression profiles of peripheral B-cells between 8 early RA patients (mean duration 1.6 years) with active disease and 8 age- and sex-matched controls (233). In RA patients there was apparent increased expression of 305 genes and reduced expression of 231 genes. Pathway analysis software identified functional clustering of the differentially expressed genes within pathways often associated with B cell function; such as, cell activation, proliferation apoptosis, autoimmunity, cytokine function and angiogenesis.

The previously described gene expression studies utilised between group comparisons of either different stages of RA or between a cohort of RA patients and an unaffected control cohort. In each case, it was inferred that the observed differences in gene expression profile were related to the phenotypic variable used to define each group (e.g. RA vs control or early RA vs established RA). All of these studies used relatively small cohort groups and therefore may not fully reflect the broad heterogeneity of RA. Only one study reported a follow-up validation analysis. A number of recent studies have taken an alternative approach whereby all participants have established diagnoses of RA and are subdivided into comparison groups using clinically relevant variables such as disease activity level or response to a specific intervention. The majority of these studies have examined the value of gene expression profiles as markers of response to biologic therapy

- Thirty three patients with DMARD resistant, established RA (mean duration 11.3 years) were treated with intravenous infliximab and provided whole blood, for PBMC gene expression profiling, immediately before, and 3 months after, commencing treatment (234). Patients were classified as being treatment responders if DAS28 fell by at least 1.2 after 3 months treatment. Differential expression analyses identified 41 gene transcripts with statistically significant ( $p = 0.05$ ) levels of expression between responders and non-responders. Quantitative Real Time-PCR reliably quantified 20 of the 41 candidate transcripts and a hierarchical clustering analysis showed that these 20 transcripts would correctly classify the treatment response in 75% of patients. Subsequent analyses identified a combination of 8 gene transcripts which were at least as accurate as the panel of 20 transcripts for classifying patient's treatment response through hierarchical clustering. Time integrated analyses comparing changes in gene expression levels between baseline and 3 months of treatment demonstrated that in responders 18 of 20 candidate gene transcripts tended towards to higher expression levels at 3 months. By contrast, in non-

responders, 19 of 20 gene transcripts exhibited a reduction in expression levels which reached statistical significance for 8 transcripts.

- Lindberg et al compared before and after treatment changes in gene expression profiles in synovial biopsy samples taken from 10 RA patients receiving intravenous infliximab (235). Patients were grouped based upon their fulfilment of EULAR response criteria after three months treatment (3 good responders, 5 moderate responders, 2 non-responders). For baseline samples, step-wise comparisons between each responder group identified 279 differentially expressed genes when good and non-responders were compared. However, there were no statistically significantly expressed genes when good responders were compared with moderate and non-responders. Following infliximab therapy, comparisons between baseline and 3 month gene expression profiles in good responders identified 115 genes whose expression levels changed significantly following therapy (i.e. a dynamic change in expression levels led to them being differentially expressed). Immunohistochemical analyses identified TNF $\alpha$  in four synovial biopsy samples, of which all were taken from either good (3) or moderate (1) responders. Comparing these samples to the TNF $\alpha$  negative biopsies identified 12 differentially expressed genes. Furthermore, comparisons between gene expression profiles for the TNF $\alpha$  positive patients with those from all non-responders identified 685 differentially expressed genes; thus suggesting that the presence of TNF $\alpha$  has pathogenic significance and may serve as an important predictor of successful infliximab treatment
- The relationship between pre-treatment gene expression profiles in synovial tissue and response to infliximab therapy has been reported in 18 patients with DMARD-resistant established RA (236). Response was defined as a reduction in DAS28 of at least 1.2 after 16 weeks of therapy. Hierarchical clustering analyses, using the 189 genes which exhibited at least a 1.4 fold between group difference in expression levels, identified a panel of genes which had clearly increased expression levels in responders and reduced levels in non-responders. These transcripts contained a number of specific genes (e.g CD163, S100A8, HLA Class II, immunoglobulin, integrins and chemokines) which have already been associated with high levels of inflammation in RA tissue
- Sekiguchi et al described how changes in serial PBMC gene expression profiles correlated to clinical changes and treatment response over time in 18 patients with DMARD-resistant RA treated with intravenous infliximab (237). Achieving an ACR50 response (i.e at least 50% improvement in core set variables) was defined as a treatment response. The investigators performed gene expression analyses using a custom made, low density (747 genes) microarray which incorporated genes known to be related to inflammatory blood cell activation. In total, 18 genes were differentially expressed (>1.5-fold change) between the responder and non-responder groups; of which, the top ten were related to interferon. In the responder group, successful treatment with infliximab caused persistent reduction in the expression levels of several interferon related genes. By contrast, in the non-responder group, infliximab therapy caused an early, transient fall in interferon-related gene

expression levels that returned to baseline levels on follow-up testing. Subsequent quantitative real-time PCR analyses confirmed the findings of the microarray analysis. Interestingly, time course analyses showed that interferon-related gene expression levels followed a very similar pattern to corresponding fluctuations in DAS28 and, in most cases, individual gene expression levels correlated strongly with corresponding DAS28 scores ( $R^2$  0.6115 – 0.8929). These preliminary results do suggest that changes in gene expression profiles may reflect changes in RA disease activity and that cut-down, customised cDNA microarrays might provide an additional means of longitudinally monitoring response to (at least) infliximab therapy.

- Nineteen patients with DMARD resistant, established RA provided whole blood PBMCs immediately before, and 3 days after, commencing treatment with etanercept to determine whether very early changes in gene expression profile were predictive of treatment response (238). Once again, a good response was defined as at least a 1.2 reduction in DAS28 3 months after commencing therapy and was achieved by 12 patients. Three days after receiving etanercept, 42 genes showed a differential change in expression levels between the responder and non-responder groups. Of these, 36 genes demonstrated were down-regulated comparing responder to non-responders. Pathway analyses demonstrated that successful etanercept therapy produced early down regulation of genes involved in pathways related to TNF $\alpha$  signalling, NF $\kappa$ B-independent signalling and regulation of cellular and oxidative stress. These early changes were associated with, and therefore may be predictive of, at least 3 month response to etanercept. Interestingly, pre-treatment gene expression profiles were not predictive of 3 month etanercept response whereas the dynamic changes in gene expression profiles after 3 days were
- The previously described study by Batliwalla et al (232) showed that there was significant correlation between disease activity levels and the expression levels of genes relating to monocyte function. Stuhlmuller et al progressed these findings by attempting to determine whether a single, monocyte-related biomarker was predictive of anti-TNF $\alpha$  response in RA (239). Monocytes were purified from blood samples donated by 77 RA patients (mean disease duration 8.7 years), who were participating in a randomised control trial of adalimumab monotherapy, and 23 healthy controls. Response to adalimumab was defined as achievement of an ACR20 response. Pairwise comparisons between RA patients and healthy controls identified 51 genes with differential expression. Hierarchical clustering analyses performed using three candidate genes correctly classified treatment response in all patients. Following adalimumab treatment, there were 117 genes which exhibited differential expression levels between responders and non-responders. Three genes (*FAM3C*, *ITGAX (CD11c)*, *TMEM45A*) were differentially expressed in all pairwise comparisons between responders and non-responders. Follow-on real time reverse transcription PCR quantification using monocytes from an independent cohort of 27 RA patients showed a strongly positive ( $r=0.651$ ), statistically significant ( $p<0.0001$ ) correlation between CD11c expression levels and ACR response to adalimumab which correctly predicted the likelihood experiencing a good treatment response. Receiver operating

characteristic analyses suggested the sensitivity for detecting responders was 100% and the specificity was 91.7%. However, the predictive utility of CD11c was only evident in patients receiving adalimumab monotherapy since it did not appear predictive of response to either methotrexate monotherapy or methotrexate/adalimumab combination therapy.

Despite their disparate designs, taken together the previously described studies do suggest that gene expression profiles might have a number of useful diagnostic and prognostic properties in relation to assessment and management of RA:

1. RA patients can be distinguished from healthy controls, and/or osteoarthritis sufferers, through specific differences in gene expression profiles (231-233,240). Furthermore, specific difference in gene expression profiles have been demonstrated between early and established RA patients (230)
2. Pre-treatment gene expression profiles correctly classify, and therefore may be predictive of, subsequent response to biologic therapy. Specifically, good clinical responses to either infliximab (234,236,237,241), adalimumab (239) or etanercept (238) have each been associated with specific pre-treatment differences in gene expression profiles between responders and non-responders
3. Treatment with biologic therapy induces a dynamic pattern change in gene expression profiles (235,237,238) which may reflect observed changes in clinical disease activity (and therefore treatment response)

To date, most transcriptomic studies in RA have attempted to compare relatively early changes in gene expression profiles to the clinical impact of a single intervention. Most commonly, this has been the introduction of biologic therapy (usually an anti-TNF $\alpha$  agent) in patients with DMARD-resistant established RA. However, as previously discussed, the time period when therapeutic intervention is most likely to have the greatest long-term benefit appears to be during the early months following symptom onset. Currently, most treatment guidelines advocate using aggressive non-biologic DMARD regimens to suppress the emerging inflammatory process and using clinical disease activity measures, with their attendant lingering concerns regarding sensitivity and specificity, to guide treatment changes. Biologic agents, the single treatment group proven to consistently retard erosive joint progression, remain reserved for those patients who demonstrate DMARD resistance through persistently active disease and, in many cases, progressive irreversible joint damage. Furthermore, the previously described transcriptomic studies are starting to identify specific gene expression patterns, and indeed differences in expression of single genes, which might in fact increase the overall efficacy of biologic therapy further by ensuring that specific therapies are targeted at those patients most likely to experience a beneficial response. However, the majority of newly diagnosed RA patients will not initially receive biologic therapy; therefore, the potential role for using gene expression profiles in the often rigidly dictated assessment and management of early RA remains to be described. If similar relationships to those demonstrated between anti-TNF $\alpha$  treatment response and gene expression

profile are also evident when applied to non-biologic DMARD treatment response it might be possible to identify specific gene expression profile patterns which either predict response or non-response to non-biologic DMARD therapy (either singularly or as a group). Ultimately, it may be possible to use gene expression information at presentation to screen newly diagnosed RA patients' blood for the presence of specific gene expression patterns to predict their likelihood of responding to a particular treatment, or regimen. In this way, specific treatments, or regimens, can be targeted at the patients most likely to respond to the treatment and, theoretically, patients should avoid the unnecessary (and uncomfortable) delays associated with receiving ineffectual therapy. Similarly, since both gene expression and disease activity levels appear to be dynamic, if strong correlations can be identified between specific disease activity states and specific gene expression patterns, it may be possible to use blood gene expression as an additional measure of global disease activity

### **1.6.2 Metabolomics**

Metabolomics uses high-throughput technologies, such as nuclear magnetic resonance spectroscopy and liquid gas chromatography, to describe the relative concentrations of metabolically active, low molecular weight compounds in individual tissues or organs (162). Evident differences in the metabolomic expression patterns between individuals with a specific illness and unaffected controls can provide important insights into the underlying disease processes. Important metabolomic signatures, with either diagnostic and/or prognostic properties, have been identified in studies of cancer (bladder, colorectal, prostate, stomach, renal, brain and lung), type 1 and type 2 diabetes mellitus, cardiovascular disease, neurological diseases (Parkinson's disease, amyotrophic lateral sclerosis, Alzheimer's disease and schizophrenia), asthma and coeliac disease (162). To date, relatively few descriptions of the diagnostic and/or prognostic properties of metabolomic profiling in RA have been published though it is feasible that phenotypic variations in RA may also exhibit characteristic metabolomic signatures.

- Madsen et al described the potential diagnostic properties of metabolomic profiling to distinguish RA patients from either psoriatic arthritis patients or healthy controls (242). Firstly, the plasma metabolomic profiles of 25 RA patients (early and established disease) were compared to 20 psoriatic arthritis patients. Combining the results for gas chromatography-mass spectrometry and liquid chromatography mass spectrometry identified 83 metabolites with differential expression levels between RA and psoriatic arthritis. A follow-on validation analysis used a subset of these differentially expressed metabolites to attempt to distinguish between 14 different RA patients and 20 healthy controls. This model correctly identified RA patients with a sensitivity of 93% and specificity of 70%. Whilst this specificity is clearly less than that offered by ACPA testing, analysing metabolomics profiles did correctly identify several patients with clinical diagnoses of RA but negative autoantibody statuses
- Huggle et al have recently reported on the diagnostic value of using metabolomic profiles to distinguish between septic and non-septic arthritis and between degenerative and inflammatory arthritides (243). Synovial fluid from 59 patients with a broad range of

rheumatological diagnoses was examined using nuclear magnetic resonance spectroscopy. Overall, synovial fluid from patients with septic arthritis demonstrated a distinctive metabolomic profile. However, there was no distinct differences observed between patients with osteoarthritis and inflammatory arthritides (including RA).

- Van Wietmarschen have described which differences occur in urinary and plasma metabolomic profiles when a patient's RA is classified according to Chinese medicine theory (244). Thirty nine RA patients completed a detailed symptoms questionnaire and were classified as having either Heat RA or Cold RA by Chinese medicine practitioners. Urinary and plasma metabolomic profiles were determined by liquid chromatography mass spectrometry. The authors report panels of 11 urinary metabolites and 8 plasma metabolites that discriminated between the Heat and Cold RA groups. Heat RA patients exhibited higher urinary levels of several metabolites related to carnitine synthesis and the authors suggest that Cold RA patients may exhibit either lower muscle mass or lesser rates of muscle breakdown. However, a clinical correlation to body habitus is not reported. Furthermore, Heat RA patients exhibited higher DHEAS levels than Cold RA patients suggesting that Cold RA patients may experience greater rates of suppression of hypothalamic-pituitary-adrenal axis function. Once again, a comparison with formal assessments of endocrine function is not reported.

Clearly, there hasn't been the same degree of investigation into the role of metabolomic signatures in the assessment of RA as has been reported for transcriptomic profiling. The available studies do seem to suggest that metabolomic profiling may at least have a role in supporting the diagnosis of seronegative RA and differentiating septic arthritis from other inflammatory arthritides. However, it is unlikely that Chinese Medicine Theory models of disease will ever be incorporated into the standard assessment portfolio of Western rheumatologists. Either way, further larger scale studies are required to better describe how metabolomic profiling performs as either 1. a diagnostic and/or prognostic tool in the assessment of patients with suspected inflammatory arthritides and 2. an additional measure of global disease activity able to accurately represent response to DMARD therapy changes.

Overall, high throughput technologies do show some promise as additional means of assessing patients with inflammatory arthritides. On the basis of available evidence it's possible that high throughput technologies (particularly transcriptomic profiling) may distinguish between different causes of inflammatory arthritis and different disease activity states and might also provide additional prognostic information regarding a patient's likely response to a particular therapy. However, these presumptions are mostly inferred from relatively small clinical studies with a heterogeneous range of study designs. Furthermore, the majority of treatment response studies have been conducted on the subset of patients receiving biologic therapy rather than those receiving non-biologic DMARDs. Such technologies will require careful, prospective validation in much larger patient cohorts to systematically describe their diagnostic, prognostic and disease activity properties before they might be incorporated into routine clinical practice. At present, most gene expression studies in RA have been conducted using either peripheral blood and/or

synovial tissue samples. In a clinical setting, the most readily available tissue for analysis remains peripheral blood and this is often collected during a patient's routine consultation. In most centres, the collection of synovial tissue is not routine, since additional facilities and training are required to perform percutaneous biopsy and many patients consider the procedure excessively invasive. If high through-put technologies were to identify a useful additional prognostic and/or disease activity profile, it will be most clinically useful if it were to be identified in peripheral blood

## 1.7 Objectives

The programme of research described in the following chapters has the following objectives:

1. To identify and recruit a cohort of patients with newly diagnosed rheumatoid arthritis to a prospective study of DMARD treatment strategy and novel disease activity assessment methods
2. To prospectively gather a broad range of clinical, laboratory and radiological outcome measures to be used to describe response to DMARD therapy at an individual and group level
3. To determine the value of adding musculoskeletal ultrasound assessment to standard clinical assessments of RA global disease activity and what impact this will have on DMARD escalation decisions. To also determine whether patients who undergo regular assessment of global disease activity by musculoskeletal ultrasound, in addition to DAS28, exhibit significantly better clinical, functional and radiological outcomes compared to those patients who undergo global disease activity assessment by DAS28 alone
4. To determine whether phenotypic variations in RA can be distinguished by distinct differences in peripheral blood gene expression profile
5. To describe the relationship between peripheral blood gene expression and response to DMARD therapy to determine whether peripheral blood gene expression profiling provides useful prognostic and/or disease activity information

## 1.8 Hypotheses

1. Regular assessment by musculoskeletal ultrasound, in addition to DAS28, will identify a higher burden of active, inflammatory joint disease than clinical assessment alone
2. Patients who undergo regular global disease activity assessment by musculoskeletal ultrasound, in addition to DAS28, will have DMARD therapy tailored more appropriately to their specific needs and experience a better benefit:risk ratio than those who undergo clinical assessment alone. Specifically, using musculoskeletal ultrasound to influence DMARD escalation decisions will produce significantly better clinical, functional and radiological outcomes than using clinical assessment alone
3. Specific RA phenotypic variations will be associated with specific peripheral blood gene expression profile signatures
4. Specific perturbations in peripheral blood gene expression profile will be associated with positive or negative treatment responses and therefore may have clinically useful predictive properties
5. Fluctuations in RA disease activity in response to DMARD therapy will be reflected by corresponding changes in peripheral blood gene expression profile. Hence, peripheral blood gene expression profiling might serve as an additional measure of global RA disease activity

## **2. Methods**

## 2.1 Introduction

The research described herein has been conducted on a single cohort of 111 patients with newly diagnosed RA who provided the necessary clinical and radiological outcome data, and additional blood samples, during monthly attendances at specially set-up rheumatology research clinics. The clinical and radiological outcome data collected as part of the musculoskeletal ultrasound research study will also be used to inform the analysis of the gene expression and multi-biomarker disease activity (MBDA) test datasets. Hence, the gene expression and MBDA analyses are considered to be nested within the main clinical study. This approach has allowed very detailed, longitudinal descriptions of individual patient's, and patient group's, response to step-up DMARD therapy at clinical, functional, radiological, gene expression and biochemical levels. The description of methods will be presented in sequence with specific subsections relating to relevant aspects of the musculoskeletal ultrasound and gene expression analyses included where appropriate.

## 2.2 Identification of Study Cohort

In order to conduct longitudinal research into early RA treatment a mechanism must exist to allow recruitment and follow-up of potential participants in a timely manner. This process must facilitate early identification and review of patients, such that the screening process used to determine whether a patient is suitable to participate in a clinical trial does not produce unacceptable delays in diagnosis or commencement of appropriate treatment. Furthermore, patients who undergo screening and either decline, or are deemed unsuitable, to participate must not experience any compromise in delivery of appropriate care. The following sections will describe how patients who participated in this research project were identified, screened for participation and the reasoning behind the final cohort size

### 2.2.1 Screening Arrangements

#### ***Research Sites***

Specific research clinics were established at three hospital sites in Glasgow (Gartnavel General Hospital, Glasgow Royal Infirmary and Stobhill Hospital) for screening and follow-up of patients. Research clinics ran 4-5 times weekly at Gartnavel General Hospital, twice weekly at Glasgow Royal Infirmary and once weekly at Stobhill Hospital. Each clinic was staffed by a Clinical Research Fellow (Dr James Dale – JD) and an experienced Rheumatology Research Nurse / Metrologist (Sister Anne Stirling – AS). The balance between new patient reviews (1 hour) and return slots (30 minutes) was varied according to the stage and recruitment status of the study. In addition to the three main hospital sites potential participants could also be referred for screening from Rheumatology Departments based at other hospital sites within Greater Glasgow and Clyde (including the Victoria Infirmary, Inverclyde Royal Hospital, Royal Alexandria Hospital, Nuffield Health Glasgow Hospital and BMI Ross Hall Hospital) with follow-up being arranged at the research clinic most convenient for the participant

## **Referral Sources**

In order to maximise the rate of recruitment, potential participants could be referred for screening by a variety of different routes:

1. Patients who had already undergone outpatient rheumatology clinic review and had recently been given presumptive diagnoses of RA were referred directly to Dr Dale by the base hospital rheumatology team. These referrals were made by either telephone, email or directly face-to-face
2. At Gartnavel General Hospital and Stobhill Hospital, new patient referrals from Primary Care or other specialities that were suggestive of RA were forwarded to Dr Dale and provided with early first review appointments at the research clinics. Patients who fulfilled inclusion criteria, and who agreed to participate, continued to attend the research clinics whereas those who either declined to participate, or did not fulfil inclusion criteria, had appropriate follow-up arranged within their base rheumatology unit
3. Acute general medical admissions with severe, debilitating first presentations of RA were referred to Dr Dale for screening and early follow-up

For each participating site, the referral of potential participants for screening was encouraged by displaying recruitment posters in outpatient clinics and day-wards, sending regular recruitment emails and updates to colleagues (consultants, trainees and specialist nurses) and by Dr Dale attending each Rheumatology Department's academic meetings to present the aims, design and treatment protocol for the study. A copy of the recruitment poster can be found in Appendix B

## **Screening Process**

All potential participants underwent a standardised screening process conducted by Dr Dale and Sister Stirling. The aims of this process were to 1. ensure that potential participants had the correct diagnosis and fulfilled the inclusion criteria, 2. ensure that potential participants neither met any of the exclusion criteria nor exhibited any other contraindication to taking part, 3. provide participants with ample opportunity to discuss their diagnosis, the nature of their treatment and the implications of taking part in clinical research and 4. collect all of the necessary baseline clinical, laboratory and radiological outcome data. During the screening process the following standardised assessments were performed:

- *Clinical*
  1. Comprehensive clinical history (JD) – including history of presenting complaint, past medical history, medication history and social circumstances
  2. Clinical examination (JD) – including general systemic examination and musculoskeletal examination
- *Laboratory*
  1. Biochemistry – Urea + electrolytes (UE), liver function tests (LFT), C-reactive Protein (CRP), calcium, glucose, non-fasting lipid profile (total cholesterol, HDL, triglycerides and cholesterol:HDL), thyroid function tests (TFT)

2. Haematology – Full blood count (FBC), Erythrocyte sedimentation rate (ESR)

3. Immunology – Rheumatoid factor (RF), anti-cyclic citrillunated peptide antibodies (anti-CCP), anti-nuclear antibody (ANA)

- Radiological
  1. Chest X-ray
  2. Plain X-ray hands and feet (AP projection)
  3. Musculoskeletal ultrasound of hands and feet (JD) – scanning of limited joint set (bilateral radiocarpal, MCP2+3, PIP 2+3 and MTP 2+5) with grading of synovial hypertrophy and PD signal
- Disease Activity
  1. DAS28 (JD) – comprising 28 swollen joint count, 28 tender joint count, patient global health 10cm visual analogue score and ESR
  2. DAS44 (AS) – comprising 44 swollen joint count, Ritchie articular index, patient global health 10cm visual analogue score and ESR
  3. Total pain 10cm visual analogue score
- Functional Assessment
  1. Health assessment questionnaire (AS)
  2. Euro-QOL 5D questionnaire (AS)

Due to the generally accepted requirement that there must elapse at least 24 hours between the initial discussion of participating in clinical research and the provision of written consent the screening process required most patients to attend for two separate visits which were usually scheduled within 7 days of each other. The short delay between screening and recruitment visit also allowed the return of any outstanding blood test results that might otherwise have precluded the patient's participation. All potential participants were provided with a standardised Patient Information Leaflet describing the nature of the project and all who agreed to participate were required to provide written consent. Examples of the Patient Information Leaflet and Patient Consent Form are reproduced in Appendices C and D. Any patients who either declined to participate, did not fulfil the inclusion criteria, or did fulfil the exclusion criteria, underwent a full clinical evaluation and - during the follow-up review - were commenced on appropriate treatment (as indicated) with onward follow-up being arranged at their base rheumatology unit.

### **2.2.2 Inclusion Criteria**

The inclusion criteria are similar to criteria used in a number of other studies of RA treatment strategy and were designed to ensure that only patients with the correct diagnosis and presenting characteristics were offered participation in the research

*Patient with newly diagnosed RA or CCP-positive Undifferentiated Inflammatory Arthritis (with 3 or more swollen joints*

Since the research aimed to be applicable to modern clinical practice a pragmatic approach to recruitment was undertaken. In order to identify a cohort of patients that was comparable to that

encountered in daily practice, patients who could justifiably be given a clinical diagnosis of RA were considered for recruitment. Furthermore, given the very high specificity of positive anti-CCP antibodies for RA (10), the evidence that inflammatory arthritis in anti-CCP positive patients has a very high risk of evolving into RA (245,246) and that anti-CCP positivity is considered a poor prognostic marker (53,247), patients with clinically undifferentiated inflammatory arthritis who were anti-CCP positive were also considered for participation. However, since anti-CCP antibodies can be detected in blood for many years before the onset of clinical inflammatory joint disease (7), and since many anti-CCP positive individuals may experience arthralgia without having evident inflammatory joint disease, anti-CCP positive patients needed to exhibit at least 3 clinically swollen joints to ensure that their symptoms could be attributed to the presence of synovitis.

#### *Symptom duration less than 12 months*

Clinical observations have demonstrated that different temporal stages of RA may react differently to immunomodulatory therapy, with the greatest chance of achieving a prolonged, beneficial response being during the very earliest stages of symptom onset before immune plasticity is lost (113,116). Thus, whilst early and established RA are on the same pathological and clinical continuum it is possible that they may require different therapeutic approaches. Hence, since the opportunity to maximise long-term treatment outcomes appears to be within the early stages after symptom onset, this research has focussed on 'early RA'.

During the period of drafting of the research protocol (April – June 2009) there was no consensus over the definition of 'early RA' based on duration of symptoms. International guidelines available at the time all stated different time intervals:

Scottish Intercollegiate Guidelines Network	symptom duration < 5 years (248)
British Society of Rheumatology	symptom duration < 2 years (168)
European League Against Rheumatism	no comment (169)
American College of Rheumatology	disease duration < 6 months (112)

Furthermore, even the recent publication of the 2010 ACR/EULAR Classification Criteria failed to clarify the definition since they stated RA should be considered whenever arthritis has persisted for more than 6 weeks but did not distinguish between early and established disease (13). For ease of use, and so as not to limit recruitment, an arbitrary symptom duration of up to 12 months was chosen for this research. Symptom duration was chosen in preference to disease duration (i.e. the time from diagnosis) to account for any prolonged delays in seeking medical review that might have resulted in some patients first presenting to the rheumatology clinic with established disease (249). Symptom duration was timed from the point that participants first experienced consistent and persistent joint symptoms (pain, swelling and/or stiffness) that, in the opinion of the clinician, were attributable to RA.

#### *Active Disease (DAS44 $\geq$ 2.4)*

A patient's diagnosis of RA is independent of their level of disease activity provided they present with an appropriate history and display supportive examination and investigation findings. At low disease activity levels (DAS44 < 2.4) the initial treatment approach may be less aggressive than in those patients with clearly more active disease. A moderate disease activity threshold for inclusion should ensure that only patients with active disease are considered for participation and that the potential benefits of receiving aggressive, rapidly escalating doses of DMARD therapy outweigh the potential risks of experiencing treatment adverse effects

#### *DMARD Naïve or DMARD Monotherapy for less than 6 weeks*

Since the research described herein focuses heavily on optimising the very early stages of DMARD therapy it is logical to only include patients at the very start of their therapy. Otherwise any participants who were already established on DMARD therapy may either be already experiencing an early benefit (thereby negatively biasing the scale of their overall treatment response) or may appear to have prolonged courses of initial DMARD monotherapy (which would be against the ethos of early tight disease control). Furthermore, since the parallel research aim is to describe changes in biomarker signatures with treatment it is logical to only consider patients in whom there is only a small chance that an additional external factor (such as DMARD therapy) might have biased the observed findings. Since some patients were referred by external rheumatology departments to the research clinics, and since the researchers were occasionally unavailable to arrange rapid reviews, up to 6 weeks DMARD monotherapy was accepted in order to maximise the recruitment rate and avoid any unacceptable delays in externally referred patients commencing appropriate therapy.

#### *Aged 18 or over*

A lower age limit of 18 years is a standard inclusion criterion in most interventional studies in rheumatology generally and RA treatment strategies generally. An upper age limit was not applied

### **2.2.3 Exclusion Criteria**

Assuming a potential participant fulfilled the inclusion criteria, the exclusion criteria were used to ensure that participants would not be placed at an increased risk of adverse effects from the treatment protocol's DMARD escalation regimen. Essentially, the exclusion criteria are common contraindications to receiving aggressive, rapidly escalating DMARD therapy - especially methotrexate which forms the crux of the DMARD regimen

#### *Significant liver disease and/or abnormality of liver function tests (baseline AST/ALT > twice upper limit of normal or alkaline phosphatase > 2.5 times upper limit of normal)*

Both methotrexate and sulfasalazine can cause aberrations of liver function that may require either dose reduction or complete cessation of therapy; hence, potentially hepatotoxic therapies are often avoided in patients with pre-existing liver disease. Further, both methotrexate and sulfasalazine undergo extensive hepatic metabolism; hence, significant liver disease can lead to the accumulation of their metabolites and an increased risk of adverse effects. Thus, patients with pre-existing liver disease might be placed at an unacceptably high risk of developing worsening

liver function through participation in the study. Furthermore, failure to utilise the proposed DMARD regimen in a subset of patients might lead to their RA appearing to be inadequately treated (within the limits imposed by their co-morbidities) and therefore could negatively bias the observed outcomes. Fluctuations in liver function test results are important indicators of evolving liver dysfunction; however, regular monitoring can prove extremely difficult if patient's liver function tests appear abnormal prior to commencing DMARD therapy

*Significant renal impairment (baseline serum creatinine > 200  $\mu\text{mol/l}$ ; eGFR < 30)*

Both methotrexate and sulfasalazine (and their derivatives) are primarily excreted in urine. Thus, significant renal failure, and a loss of excretory renal function, can potentially lead to the accumulation of non-excreted metabolites and an increased risk of adverse effects

*Significant cytopenia (baseline white cell count < 4.0 x 10<sup>9</sup>/l; haemoglobin < 10 g/l, platelet < 150 x 10<sup>9</sup>/l)*

Both methotrexate and sulfasalazine can cause suppression of blood cell counts, through either direct myelotoxicity or an anti-metabolite effect. These effects are often unpredictable and can relate to either an isolated cell line or whole blood. In either case, exacerbating a pre-existing cytopenia would place the affected patient at an even greater risk of becoming symptomatic (e.g. worsening leucopenia = risk of opportunistic infection; worsening anaemia = risk of constitutional symptoms and cardiorespiratory compromise; worsening thrombocytopenia = risk of spontaneous or uncontrolled haemorrhage). DMARD monitoring relies upon a change in full blood count parameters to signal the development of possible DMARD associated adverse effects. However, accurate monitoring can become difficult, and potentially unsafe, when patients have pre-existing full blood count abnormalities that might prevent any additional abnormalities being highlighted.

*Pregnancy, planned pregnancy or breast feeding*

Use of methotrexate whilst pregnant leads to a high risk of either birth defects or spontaneous termination. Furthermore, low concentrations of methotrexate are excreted in breast milk and could feasibly be consumed by a suckling infant. Thus, the Federal Drug Agency (FDA) (pregnancy category X) and American Academy of Paediatricians consider methotrexate use to be strongly contraindicated during pregnancy and breast feeding. Since methotrexate forms the crux of this research's DMARD treatment regimen it would not be appropriate to offer participation to any patient (who otherwise meets inclusion criteria) who is either pregnant, currently breastfeeding or considering pregnancy since it would not be possible to fully investigate the impact of the intervention under investigation and could place the participant and their child at an unacceptably high risk of harm. Sulfasalazine use during pregnancy appears to be safe (FDA pregnancy category B) and only very small concentrations appear to be excreted in breast milk. The FDA have not formally assigned hydroxychloroquine to a pregnancy category. Cohort data suggests the risk of exposing unborn foetuses to hydroxychloroquine is equivalent to the risk of foetuses born in mothers with similar medical conditions who do not take hydroxychloroquine. A very small concentration of hydroxychloroquine is excreted in breast milk. To date, no consistent, significant evidence of harm to the infant has been demonstrated and overall the benefits of breast feeding are felt to outweigh the risk associated with hydroxychloroquine. The FDA have classified etanercept as pregnancy category B; however, there is very little data available from

human pregnancy studies. There is little longitudinal data to describe the potential risk of etanercept during breastfeeding. Since it has a high molecular weight and is not orally absorbed, only very small quantities will be expressed in breast milk and it is unlikely to affect the suckling infant. Regardless, until more detailed safety data is available the manufacturers (Pfizer) advise avoiding use of etanercept during pregnancy and breastfeeding.

#### *Contraindication to MRI*

Magnetic resonance imaging (MRI) of dominant wrist and MCP joints is the primary radiological outcome measure. However, since MRI uses a high energy magnetic field scanning is contraindicated in any patients whose body contains any ferromagnetic object for fear that the object may cause local trauma through becoming dislodged or heated by induction. If a potential participant were unable to undergo MRI they would not be able to contribute to one of the research's main outcome measures and therefore their omission could bias the observed results. Common contraindications to MRI include: implantable cardiac pacemakers and defibrillators, vagus nerve stimulators, cochlear implants, deep brain stimulators, particular cerebral aneurysm clips, particular surgical prostheses and retained metal fragments. The MRI scanner usually requires patients to lie prone in a narrow, horizontal corridor with most of their upper body surrounded by the scanning machinery. Some patients find this a very unpleasant and oppressive experience; hence, severe claustrophobia is also considered a relative contraindication to MRI

#### *Other co-morbid condition that in the opinion of the investigator would preclude the use of combination DMARD therapy*

Whilst it was hoped that the previously described exclusion criteria would capture most of the common contraindications to participating in the research it was acknowledged that there would be a small number of patients who did not fulfil the exclusion criteria in whom participation would still be considered inappropriate because of additional external factors. Thus, this final exclusion criterion was included to allow the researchers a degree of clinical judgement when considering potential participants with either multiple medical co-morbidities or significant psychological and/or social difficulties

### **2.2.4 Sample Size Estimation**

A sample size calculation was conducted to ensure adequate power to demonstrate a statistically significant difference in the primary clinical outcomes between the intervention groups in the treatment strategy arm of this research. The previously described TICORA (118) and TEAR (146) studies had been conducted in a similar geographical population with similar designs. Both of these studies had performed sample size calculations to ensure sufficient patients were recruited to demonstrate a statistically significant between-group difference. The parameters that formed the basis of these calculations are described in Table 2

	Comparator	Minimum Important Change	Standard deviation	Alpha value	Power	Group Size
<b>TICORA</b>	Mean fall in DAS44  Percentage EULAR Good	$\Delta\text{DAS44} \geq 1.1$	0.7	$p < 0.01$	95%	53
<b>TEAR</b>	Mean fall in DAS28	$\Delta\text{DAS28} \geq 1.1$	1.46	$p < 0.05$	90%	37

**Table 2** – Summary of sample size calculation parameters for TICORA and TEAR studies

In contrast to the TICORA and TEAR studies, this study used mean change in DAS44 and MRI RAMRIS scoring as co-primary outcome measures. At the time of study design available published data had shown the standard deviation of RAMRIS synovitis score at the wrist to be  $\pm 1.64$  (250) but there were no published values for the minimum clinically important difference. Presuming that the sample size estimate for the TICORA study DAS44 outcome would still hold true the following calculation was performed to infer what difference in RAMRIS synovitis scores it might be able to detect using similarly sized groups:

$\text{If } n \text{ per group} = \frac{2 \times \sigma^2 \times f(\alpha, \beta)}{(\mu_2 - \mu_1)^2}$	<p>Where <math>(\mu_2 - \mu_1)</math> = minimum difference  <math>\sigma</math> = standard deviation = 1.64  <math>\alpha</math> = probability of type 1 error = 0.05  <math>\beta</math> = probability of type 2 error = 0.2  <math>f(\alpha, \beta) = 7.9</math></p>
$53 = \frac{2 \times (1.64)^2 \times 7.9}{(\mu_2 - \mu_1)^2}$	
$(\mu_2 - \mu_1)^2 = \frac{2 \times (1.64)^2 \times 7.9}{53} \rightarrow (\mu_2 - \mu_1) = 0.90$	

Thus, groups comprising at least 53 patients would be able to detect a minimum change in RAMRIS synovitis score of 0.90 at power of 80% and alpha value  $p < 0.05$ .

## 2.3 DMARD Treatment Protocol

Since the interventional component of the TaSER study aimed to be clinically relevant a DMARD escalation protocol was devised that closely reflected current clinical practice and was supported by a sound evidence base. Patients in both treatment groups would progress through exactly the same treatment steps if their measured disease activity levels were greater than the escalation threshold. However, the method of global disease activity assessment used, and thus the threshold for DMARD escalation, would be clearly different

### 2.3.1 Overview of DMARD Escalation Protocol

A previously published description of the prescribing habits of UK rheumatologists demonstrated that the vast majority (97%) prefer to commence newly diagnosed RA patients on DMARD monotherapy and 'step-up' to combination DMARD therapy if measures of disease activity fail to improve (136). This step-up DMARD escalation approach forms the basis of practice in the majority of NHSGGC early arthritis clinics. Indeed, several treatment strategy studies have

demonstrated that early step-up DMARD therapy can produce short term clinical outcomes which are at least equivalent to more aggressive forms of initial combination therapy (123,145,146,251) and that short term delays in commencement of combination therapy, which allow a trial of DMARD monotherapy, do not excessively disadvantage clinical and radiological outcomes(148,252). This step-up approach has the additional advantage of preventing those good prognosis patients who respond adequately to initial monotherapy having to take unnecessary medications that in some cases can be extremely complicated and potentially toxic. The evidence base behind the DMARD escalation protocol is discussed in detail in Section 2.3.3

### ***General Principles***

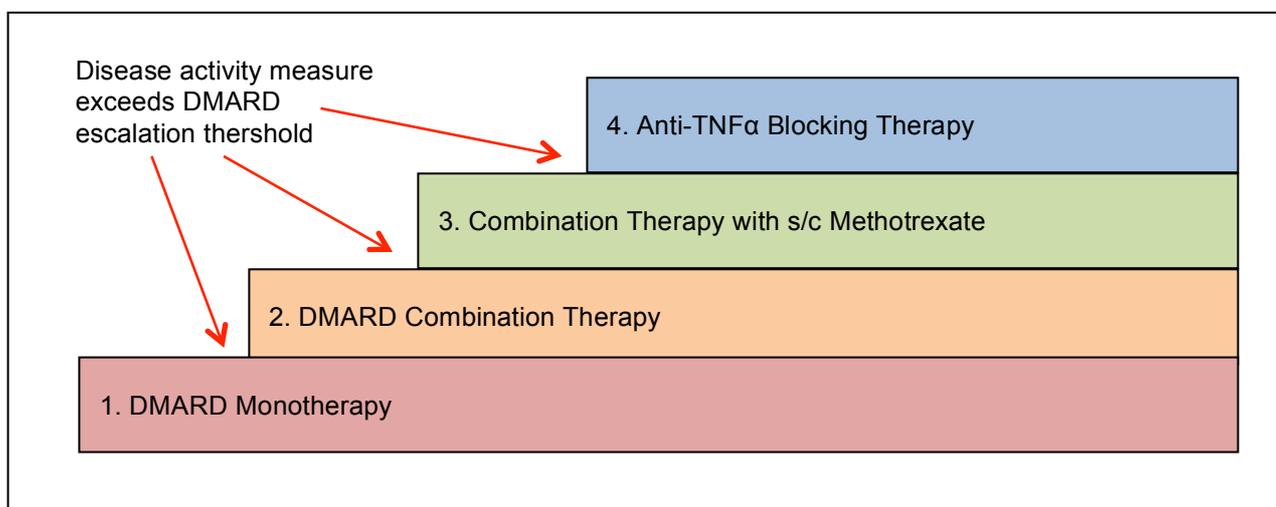
- Patients will attend for monthly clinical reviews for a total duration of 18 months
- During each review, implementation of some or all of the following aspects of a complex therapeutic intervention will be considered:
  1. Optimisation of DMARD therapy dose
  2. Administration of intra-articular and/or intramuscular corticosteroids
  3. Prescription and optimisation of appropriate level of analgesia: including NSAIDs, opioid and non-opioid analgesics and adjuvant analgesics
  4. Provision of appropriate orthoses and joint support devices
  5. Referral for assessment and treatment by members of the AHP multidisciplinary team. All patients undergoing screening will be referred for initial review by the physiotherapists and occupational therapists. During the study follow-up period it may become necessary to arrange referral to some, or all, of the following AHPs: physiotherapists, occupational therapists, podiatrists and/or dieticians
  6. Referral for assessment by members of the Orthopaedic Surgery team (as necessary)
  7. Referral for assessment and treatment of other co-morbidities by clinicians within other specialities (as necessary)

Participants in the control and intervention groups followed the same DMARD escalation steps and were offered all other aspects of the complex therapeutic intervention without bias.

- DMARD therapy was escalated until an individual participant's assigned global disease activity measure fell, and remained, below a predefined lower disease activity target (discussed in more detail in following sections)
- The effectiveness of individual DMARDs was optimised by attempting to increase each agent to either the optimum dose for each patient's weight or the maximally tolerated dose not associated with adverse effects
- Participant's DAS28 was assessed at every monthly visit. Decisions to escalate DMARD therapy were deferred until at least 3 months had elapsed following transition to each

DMARD step. This duration provided sufficient time for the full impact of the new DMARD regimen to become fully apparent

- Participants attended for monthly clinical appointments throughout the follow-up period. This frequency allowed optimisation of DMARD doses, early intervention for adverse effects and frequent administration of intra-articular and/or intra-muscular corticosteroid injections in participants who continued to exhibit active inflammatory disease. In the 3 months following progression to each DMARD escalation step, doses were optimised but no new DMARD agents were added
- This research aimed to be directly applicable to routine clinical practice and reflect the practicalities of implementing an aggressive DMARD escalation protocol. Therefore, participants who failed to tolerate an individual agent, or who declined DMARD escalation even if indicated by the study protocol, continued to be followed up via the research clinics and were not deemed treatment failures
- The metrologist (AS) scoring the clinical, functional and radiological outcome data was blinded to each participant's treatment group for the duration of the research. Dr Dale, who was responsible for the clinical assessment and management of each patient, was not blinded to participant's treatment group since he was required to perform the relevant clinical and MSUS assessments of global disease activity for each participant and to decide when DMARD escalation was indicated. The collection of any blinded outcome data was not used to influence any decision to escalate DMARD therapy



**Figure 4:** Diagrammatic representation of DMARD escalation protocol

## ***DMARD Escalation Steps***

Figure 4 provides a diagrammatic representation of the DMARD escalation protocol.

### *Step 1 – DMARD Monotherapy*

Methotrexate 7.5mg/wk with folic acid 5mg/wk  
Increased monthly to target dose - 7.5 mg/wk → 15 mg/wk → 20 mg/wk

OR

Sulfasalazine 500mg/d  
Increased weekly to target dose (approximately 400mcg/kg/d)

### *Step 2 – Combination DMARD Therapy*

Triple therapy Methotrexate, sulfasalazine and hydroxychloroquine

OR

Dual therapy Patients who are intolerant of a single agent will receive the other two agents in combination

### *Step 3 – Combination Therapy with Parenteral Methotrexate*

Conversion of oral methotrexate to equivalent subcutaneous dose. If oral methotrexate escalation was limited by intolerance, attempts will be made to escalate subcutaneous methotrexate dose to maximum (25mg/wk) or maximum tolerated dose

### *Step 4 - Anti-TNF $\alpha$ Therapy*

Addition of subcutaneous etanercept (50mg/wk) to combination DMARD therapy

### *Step 5 - Anti-TNF $\alpha$ Therapy Withdrawal*

Patients who achieve remission with etanercept (Control group - DAS28<2.6, MSUS group – no PD signal) at three consecutive monthly assessments will have etanercept discontinued after 6 months of treatment. In event of a clinically evident acute flare etanercept will be restarted indefinitely

## ***Corticosteroid Therapy***

Corticosteroid therapy has a rapid onset, immunomodulatory effect which can often provide patients with short-term symptomatic respite whilst longer, and slower, acting DMARD therapy is being established. Multiple studies have attempted to describe the potential benefits of corticosteroids as treatment for RA; however, in many cases the results are often contradictory. Eitherway, treatment guidelines consistently recommend that corticosteroids in some form be administered in combination with DMARD therapy in patients with early RA (111,112,124,169,170). Theoretically, the rapid immunomodulatory effect of corticosteroids should produce profound, early improvements in measures of inflammatory disease activity which will serve to restrict an individual patient's overall cumulative inflammatory burden. In the management of early RA, corticosteroid therapy has been shown to produce the following improvements in outcome:

1. Rapid anti-inflammatory effect. Significantly greater improvements in measures of disease activity and higher remission rates have been demonstrated when corticosteroids are given concomitantly with DMARDs compared to DMARD monotherapy (253) (254) (255)
2. Short-term disease modifying effects. Initial co-prescriptions of corticosteroids and DMARDs are consistently associated with reduced rates of radiographic progression compared to DMARD therapy alone. (254,256-258)
3. Stable disease remission. Patients treated with corticosteroids and DMARD combinations are significantly more likely to achieve, and maintain, clinical remission than those receiving DMARD therapy alone (259)

Within NHSGGC rheumatology departments corticosteroids have traditionally been administered either parenterally (particularly intra-muscularly) or as direct intra-articular injections in isolated joints. This pragmatic practice ensures patients still experience the rapid initial immunomodulatory effects of corticosteroids, whilst limiting the overall systemic dose and avoiding the need for complicated drug regimens comprising multiple tablets and frequent dose changes. Indeed, a recent study of treatment strategy has suggested that short term clinical outcomes in early RA are equivalent in patients prescribed either oral or intramuscular corticosteroids (260). Until disease activity levels fall consistently below DMARD escalation thresholds this studies treatment protocol will aim to actively administer corticosteroid therapy in addition to the escalating DMARD regimen. The corticosteroid treatment offered will be guided by the following principles:

1. Corticosteroid therapy to be considered whenever clinically active, inflammatory disease is evident
2. Corticosteroid therapy to be preferentially administered as either intra-muscular and/or intra-articular injections of triamcinolone acetonide (or equivalent) and up to 120mg given at each clinical review
3. Intra-articular steroid injections to be considered for any clinically swollen joint if that joint has not been injected within the preceding 3 months
4. Short, tapering courses of oral corticosteroids to be reserved for any patients who demonstrates persistently active moderate-high disease activity despite multiple attempts at intra-muscular and/or intra-articular steroid injection

### **2.3.2 Clinical Factors Relating to Individual Agents**

The following section describe practical and clinical issues relating to the individual DMARDs that make up this research's DMARD escalation regimen. The individual dosing schedules for each agent were based on those employed by NHSGGC rheumatology departments at the time the protocol was devised. The monitoring requirements are broadly in keeping with those suggested by the British Society of Rheumatology (261). In each case, attempts were made to escalate new DMARDs in the stages described below to either the target, or highest tolerated, dose. In many

cases intervening clinical factors – such as intolerance or adverse effects – lead to a less than target dose being accepted long term

**Methotrexate** – is an antimetabolite. Structurally, it is a weak dicarboxylic acid which closely resembles, and therefore competitively inhibits, dihydrofolic acid

**Mechanism** – the exact method by which methotrexate exerts its immunomodulatory effect in RA is not known. Methotrexate inhibits folic acid metabolism through directly inhibiting several intracellular enzymes, most notably dihydrofolate reductase. This leads to reduction in purine and pyrimidine synthesis, reduced nucleic acid synthesis, inhibition of inflammatory cytokine production and cellular adhesion molecule expression, promotion of cell apoptosis and extracellular adenosine release (262). Ultimately, there is suppression of T-lymphocyte activation and proliferation

#### *Dosing and escalation regimen*

Month 1	MTX 7.5mg/wk with folic acid 5mg/wk
Month 2	MTX 15mg/wk with folic acid 5mg/wk
Month 3+	MTX 20mg.wk with folic acid 5mg/wk

**Monitoring requirements** - fortnightly FBC, U+E and LFT until both methotrexate dose and blood monitoring have been stable for at least 6 weeks; thereafter, monthly monitoring of FBC, U+E and LFT

**Adverse Effects** – nausea, abdominal pain and anorexia (especially post-dose), ulcerative stomatitis and mouth ulcers, fatigue, macrocytosis and cytopenias (including leucopenia, anaemia, thrombocytopenia or pancytopenia), acute hepatitis, acute pneumonitis and pulmonary fibrosis

#### *Cautions*

- i. Alcohol intake – restricted to 4-6 units per week to reduce risk of hepatotoxicity
- ii. Drug interactions – increase serum methotrexate concentration and therefore increase risk of toxicity; co-trimoxazole, trimethoprim, phenytoin, probenecid, tolbutamide, NSAID (especially diclofenac)
- iii. Renal impairment – significant renal impairment will restrict methotrexate excretion and therefore increase the risk of toxicity

**Sulfasalazine** - comprises sulfapyridine and 5-aminosalicylic acid joined by an azo bond.

**Mechanism** – sulfasalazine's immunomodulatory mode of action is not clearly understood. Following ingestion, sulfapyridine and 5-aminosalicylic acid are separated by intestinal bacteria. 5-aminosalicylic acid is considered the active component. Sulfasalazine suppresses folic acid metabolism through the inhibition of 5-aminoimidazole-4-carboxamide ribonucleotide transformylase (263), promotes extra-cellular adenosine release (264) and inhibits release of pro-inflammatory cytokines (265)

*Dose and escalation regimen* – daily dose is increased every week in 500mg steps until daily target dose (approximately 400mcg/kg/d) is achieved.

*Monitoring requirements* - monthly FBC and LFTs for 3 months; thereafter 3 monthly

*Adverse Effects* – nausea, vomiting, abdominal pain, anorexia, diarrhoea, headache, rash, oral ulceration, macrocytosis and cytopenia (leucopenia, anaemia, thrombocytopenia and/or pancytopenia) acute hepatitis

#### *Cautions*

- i. Sulphonamide hypersensitivity
- ii. Glucose-6-phosphate dehydrogenase deficiency – increase risk of haemolysis
- iii. Renal impairment – increase risk of crystalluria

**Hydroxychloroquine** – is an anti-malarial agent which has also been used to reduce inflammation in a variety of inflammatory disorders

*Mechanism* – hydroxychloroquine accumulates within intra-cellular lysosomes to cause significant increases in inter-lysosomal pH, a reduction in proteolysis and a reduction in secretion of pro-inflammatory mediators. Furthermore, hydroxychloroquine restricts the activation of pro-inflammatory lymphocytes through interfering with antigen presentation by MHC class II proteins (266)

*Dose and escalation regimen* – dosing is based upon the patient's weight (upto 6.5mg/kg). There is no dose escalation schedule

Weight ≤ 46kg	hydroxychloroquine 200mg/d
46 kg < Weight ≤ 62kg	hydroxychloroquine 300mg/d (400 / 200mg alternate days)
Weight > 62 kg	hydroxychloroquine 400mg/d

*Monitoring requirements* – specific blood monitoring is not required. The Royal College of Ophthalmologists recommend pre-treatment visual screening and annual formal assessment of visual acuity using standardised reading charts. Referral for an ophthalmologist's opinion is indicated if either patients: 1. are unable to complete baseline visual screening or 2. develop worsening visual acuity whilst prescribed hydroxychloroquine (267)

*Adverse Effects* – rash, pruritus, anorexia, nausea, vomiting, diarrhoea, lethargy, cytopenias (leucopenia, anaemia, thrombocytopenia and/or pancytopenia), irreversible retinal toxicity, skeletal and cardiac muscle myopathy

#### *Cautions*

- i. Epilepsy – hydroxychloroquine may lower seizure thresholds
- ii. Psoriasis – hydroxychloroquine may cause acute exacerbations of skin plaques
- iii. Drug interactions – hydroxychloroquine may cause increased serum concentrations of digoxin, methotrexate and ciclosporin

**Etanercept** – is a fusion protein produced by DNA engineering which links the gene coding for soluble TNF $\alpha$  receptor 2 to the gene coding for the Fc component of human immunoglobulin G1 (268).

**Mechanism** – etanercept suppresses active inflammation by acting as a decoy receptor that binds and neutralises circulating TNF $\alpha$ . This mimics the action of naturally occurring, soluble TNF $\alpha$  receptors. TNF $\alpha$  is a potent, pro-inflammatory cytokine that promotes the proliferation, differentiation and migration of pro-inflammatory cells into areas of active inflammation. Inhibition of TNF $\alpha$  activity restricts the influence of a major positive feed back loop that would otherwise perpetuate inflammation

**Dose** – etanercept 50mg/wk subcutaneously for 6 months, usually self-administered by the patient. There is no dose escalation and all other non-biologic DMARDs are continued without dose adjustment. It is preferable for patients to remain on methotrexate throughout the course of etanercept

**Monitoring requirements** – etanercept is normally co-prescribed with non-biologic DMARDs, therefore the monitoring of the non-biologic DMARDs will take precedence over, and account for, any monitoring requirements for the etanercept

**Adverse Effects** – injection site reactions, pruritis, allergic reaction, increased risk of typical (e.g. upper respiratory tract, skin and soft tissue, urinary tract) and atypical (e.g. tuberculosis) infection, reactivation of latent infection (especially tuberculosis), psoriaform rash, lupus-like syndrome, leucopenia, demyelination

**Pre-treatment Screening** – due to the increased risk of infection, tuberculosis reactivation and demyelination in patients treated with any of the anti-TNF $\alpha$  agents, all patients who fulfil criteria for etanercept will be subjected to the same pre-treatment screening procedures as all other NHSGGC patients who are considered for anti-TNF $\alpha$  blocking therapy

- i. Safety questionnaire – including tuberculosis risk factor assessment, risks for intercurrent infection (indwelling catheter, previous history of septic arthritis), family history of demyelinating illness and previous history of cancer
- ii. Chest X-ray – for evidence of previous tuberculosis exposure
- iii. Blood T-spot Test – interferon-gamma release assay which detects the presence of effector T-lymphocytes sensitised against *Mycobacterium tuberculosis* and therefore identifies patients with previous, currently inactive, tuberculosis (i.e. latent disease)

If any of the methods used during pre-treatment screening identify that a participant has either been previously exposed to tuberculosis, or has evidence of latent disease, a 3-month course of

isoniazid chemoprophylaxis will be commenced prior to commencement of treatment with etanercept

### **2.3.3 The Evidence Supporting the DMARD Escalation Regimen**

The following sections will summarise the evidence base behind the use of each DMARD agent and the sequence of escalation steps used within the treatment protocol

#### **Step 1 DMARD Monotherapy**

Most of the data comparing the head-to-head effectiveness of individual DMARDs is limited to single studies where there will often be variability in the structure and quality of study design. A systematic review of published DMARD strategies has suggested that, in head-to-head comparisons, there is no significant difference in the efficacy of either methotrexate, sulfasalazine or leflunomide (137). Within NHSGGC Rheumatology departments, leflunomide is not a commonly used initial DMARD and there is less evidence available to support its use in combination with other DMARDs. Thus for the purposes of this research leflunomide was excluded from the DMARD protocol. Two randomised studies have compared the efficacies of methotrexate and sulfasalazine head-to-head and a large cohort study has described the impact of both agents on all cause mortality:

- Haagsma et al randomised 105 patients with untreated early RA (symptom duration < 12 months) to receive treatment with either sulfasalazine (SASP) monotherapy, methotrexate (MTX) monotherapy or both in combination (COMB) (141). After 1 year's follow-up, there were no statistically significant differences in clinical outcomes between any of the treatment groups. Similar mean changes in DAS44 (SASP -1.6 vs MTX -1.7 vs COMB -1.9), fulfilment of ACR response criteria (SASP 25 vs MTX 25 vs COMB 28) and fulfilment of EULAR good criteria (SASP 14 vs MTX 15 vs COMB 14) were reported. Furthermore, overall adverse event rates were similar between groups though a higher number of patients treated with sulfasalazine withdrew because of adverse events. Radiographic progression scores were not reported
- Dougadas et al randomised 209 patients with untreated, active (DAS44 > 3.0) seropositive early (disease duration < 12 months) RA to receive either sulfasalazine monotherapy, methotrexate monotherapy or both agents in combination (142). Overall, patients in all treatment groups experienced similar improvements in most disease activity measures; however, patients receiving combination therapy did exhibit a significantly greater mean improvement in Ritchie Articular Index (SASP -7.1 vs MTX -4.2 vs COMB -9.4;  $p = 0.001$ ) and a significantly better improvement in DAS44 (SASP -1.15 vs MTX -0.87 vs COMB -1.26;  $p = 0.019$ ). ACR and EULAR good response rates were similar across the three treatment groups (ACR: SASP 59% vs MTX 59% vs COMB 65%; EULAR good SASP 34% vs MTX 38% vs COMB 38%). Compared to methotrexate monotherapy, patients treated with sulfasalazine monotherapy tended to experience slightly greater mean improvements in Ritchie Articular Index (-7.1 vs -4.2), 44 swollen joint count (-4.5 vs -3.9), ESR (-30 vs -24) and DAS44 (-1.15 vs -0.87). Whilst methotrexate monotherapy patients tended to

exhibit slightly greater mean improvements in duration of morning stiffness (-53 minutes vs -46 minutes) and CRP (-16 vs -8). However, statistical comparisons between the monotherapy group outcomes are not reported. All three treatment groups demonstrated comparable rates of progression in measures of radiographic progression (mean change modified Total Sharp Score: SASP 4.64 vs MTX 4.50 vs COMB 3.46;  $p>0.05$ ). Total adverse event rates were higher in the monotherapy groups but significantly higher in the combination therapy group (SASP 75% vs MTX 75% vs COMB 91%;  $p=0.025$ ).

- Choi et al have described the impact of different DMARDs on all cause mortality in a large cohort ( $n=1240$ ) of RA patients who have been prospectively followed-up from the time of diagnosis (269). Over the follow-up period, 588 patients had ever received methotrexate, 191 patients died, of whom 72 had received methotrexate. Cardiovascular disease was the most common cause of death (44%). Even correcting for the presence of poor RA prognostic factors, methotrexate use was associated with a significant reduction in the risk of all cause mortality (hazard ratio 0.4 (CI 0.1-0.7)) and the risk of cardiovascular mortality (hazard ratio 0.3 (CI 0.2-0.7)). Furthermore, this positive impact on mortality outcomes was not demonstrated with any another DMARD (hazard ratios: MTX 0.2 vs SASP 0.9 vs penicillamine 0.8 vs hydroxychloroquine 0.7 vs IM gold 1.9)

Taken together these results do seem to suggest that, when used as DMARD monotherapy, there is little to separate the clinical efficacy of methotrexate and sulfasalazine since they are both associated with the same degree of improvement in clinical outcome measures and similar rates of radiographic progression. However, methotrexate does appear to be better tolerated than sulfasalazine and has a clear influence of mortality outcomes that is not apparent with sulfasalazine. Thus, whilst patients participating in this study will be offered either methotrexate or sulfasalazine as initial DMARD monotherapy, the adverse effects profiles and mortality benefits lend a strong weighting towards methotrexate.

### ***Step 2 Combination DMARD Therapy***

Several studies have demonstrated that adding either ciclosporin (132), leflunomide (134) or parenteral gold (135) to established methotrexate monotherapy can produce additional improvements in clinical outcomes. However, these are combinations that are often associated with adverse events, are not commonly used in local practice. Given the potential number of drugs that have either proven disease modifying properties, or have traditionally been considered DMARDs, there are theoretically a vast number of DMARD combinations possible. Most of the recent RA treatment strategy studies have focussed on triple DMARDs combinations that incorporate methotrexate (MTX), sulfasalazine (SASP) and hydroxychloroquine in varying sequences and doses. Table 3 summarises the differences in use of triple DMARD therapy in these studies. Importantly, studies of DMARD escalation strategies following methotrexate failure have not shown any inferior clinical outcomes when stepping-up to triple DMARD therapy compared to stepping-up to methotrexate and biologic (123,149). In particular, patients in Group 2 (step-up combination therapy) of the BeST study demonstrated equivalent improvements in the primary outcome measure (change in HAQ) compared to those in Group 4 who were initially

treated with biologic therapy (145). There were some advantages to biologic therapy since a greater proportion of Group 4 patients did achieve persistent low disease activity (Group 2 = 21% vs Group 4 = 40%) and exhibited lesser rates of radiographic progression (mean change total Sharp score: Group 2 5.2 vs Group 4 2.5; mean change in erosion score: Group 2 3.1 vs Group 4 1.3). In the SWEFOT study, patients who stepped-up to triple DMARD therapy after methotrexate failure exhibited statistically similar ACR (all levels) and EULAR Good response rates compared to those stepped-up to methotrexate and infliximab (147). Both groups exhibited evidence of radiographic progression but there was no statistically significant difference in the degree of progression observed between either group.

	<b>Step 1</b>	<b>Step 2</b>	<b>Step 3</b>
<b>TICORA</b> (118)	SASP	MTX – SASP – HCQ	
<b>TEAR (UK) – Control</b> (146)	SASP	MTX – SASP - HCQ	
<b>TEAR (UK) – Triple</b>	MTX – SASP - HCQ	Optimise MTX dose	
<b>BEST - Group 2</b> (145)	MTX	MTX - SASP	MTX – SASP – HCQ
<b>SWEFOT – Group A</b> (147)	MTX	MTX – SASP - HCQ	

**Table 3** – Summary of DMARD escalation steps in early RA treatment strategy studies

Several, older, randomised controlled trials, which were not necessarily conducted in early RA, have demonstrated that triple DMARD therapy tends to be more effective than any combination of dual DMARD therapy.

- O’Dell et al randomised 102 patients with active RA, despite DMARD monotherapy to switch to either methotrexate monotherapy, sulfasalazine-hydroxychloroquine dual therapy or triple therapy with all three agents (270). 50 patients fulfilled the primary outcome criteria by achieving at least 50 percent improvement in disease activity measures after 9 months and sustaining it over 2 years of follow-up. A statistically significant greater proportion of patients receiving triple therapy (77%) achieved the primary outcome compared to those receiving either dual therapy (40%) or methotrexate monotherapy (33%). Furthermore, patients receiving triple therapy exhibited lower tender joint counts (mono 7 vs dual 7 vs triple 3), swollen joint counts (mono 5 vs dual 7 vs triple 2) and ESR (mono 16 vs dual 16 vs triple 10) but not an increased rate of adverse effects.
- Calguneri et al randomised 180 patients with active, untreated RA to receive treatment with either monotherapy (MTX or SASP or HCQ), dual therapy (MTX-SASP or MTX-HCQ) or triple therapy (MTX-SASP-HCQ) (271). Results for each group were pooled rather than being reported for each of the potential combinations. After 2 years therapy, patients in all groups experienced significant improvements in clinical and laboratory outcome measures. The magnitude of response observed was related to the intensity of the DMARD regimen with patients receiving triple therapy exhibiting significantly greater responses than the dual therapy group who exhibited significantly greater responses than the monotherapy group

(Improvement >50%: mono 49.1% vs dual 73.2% vs triple 87.9;  $p < 0.001$  for all comparisons. ACR remission rates: mono 31.5% vs dual 44.6% vs triple 60.3%;  $p = 0.007$  for all between group comparisons). Rates of radiographic non-progression were significantly higher in both combination therapy groups (mono 24.5% vs dual 64.2% vs triple 68.9)

- 171 patients with active RA despite DMARD monotherapy were randomised to receive either methotrexate-sulfasalazine dual therapy, methotrexate-hydroxychloroquine dual therapy or methotrexate-sulfasalazine-hydroxychloroquine triple therapy (131). After 2 years follow-up, ACR20 response rates were significantly higher for patients receiving triple therapy (78%) compared to patients receiving either methotrexate-hydroxychloroquine (60%;  $p = 0.05$ ) or methotrexate-sulfasalazine (49%;  $p = 0.002$ ) dual therapy. Triple therapy produced significantly higher ACR50 response rates (55%) compared to methotrexate-sulfasalazine (29%;  $p = 0.005$ ) but not methotrexate-hydroxychloroquine (40%;  $p = 0.10$ ). Furthermore, there were no statistical between group differences for the ACR70 response rates (26% vs 18% vs 16%). Drug toxicity rates were similar across all treatment groups. Radiological progression rates were not reported

Thus, this study has shown that patients who exceed DMARD escalation thresholds after at least 3 months of DMARD monotherapy, will step-up to triple DMARD therapy, since this appears to be produce greater improvements in clinical outcomes compared to any of the available dual therapy combinations. Whilst combining multiple DMARDs with multiple modes of action should theoretically improve the likelihood of achieving a beneficial treatment response, it could also increase the risk of drug associated adverse effects. Individual DMARD combination therapy trials have tended not to demonstrate an increased risk of adverse effects compared to DMARD monotherapy with the constituent agents. A large meta-analysis of 36 DMARD strategy studies concluded that whilst combination therapy per se appeared more effective than monotherapy (RR 0.35; 95%CI 0.28, 0.45;  $p = 0.00001$ ) there was also a higher risk of toxicity (RR 1.37; 95%CI 1.16, 1.62;  $p = 0.0001$ ) (144). However, importantly, combinations incorporating methotrexate, sulfasalazine and/or hydroxychloroquine were not shown to have a greater risk of toxicity compared to monotherapy (RR 0.81; 95%CI 0.52, 1.27;  $p = 0.66$ )

### ***Step 3 Combination Therapy with Subcutaneous Methotrexate***

Current British Society of Rheumatology Guidelines require RA patients to have tried 2 or more non-biologic DMARDs, for at least 6 months each, before they can be offered anti-TNF $\alpha$  blocking therapy (272). The DMARD protocol for this study aims to aggressively escalate DMARD therapy every 3 months if active disease persists. Thus, if followed in their entirety, the BSR anti-TNF $\alpha$  prescribing criteria, could potentially leave patients with persistently active disease 'waiting' for 6 months to elapse following escalation to triple DMARD therapy until they became eligible for biologic therapy. This delay is contrary to the principle of tight, early disease control and minimisation of cumulative exposure to inflammatory disease that this study is hoping to optimise. If DMARD escalation thresholds continue to be exceeded after at least 3 months of triple DMARD therapy, switching from oral to subcutaneous methotrexate may allow optimisation of the efficacy

of the DMARD therapy whilst also satisfying the BSR guidelines for duration of DMARD therapy (273).

Pharmokinetic studies have demonstrated that low dose oral methotrexate preparations have a highly variable bioavailability, ranging between 25% to 100% of the ingested dose (274). In contrast, the bioavailability of parenteral methotrexate (subcutaneous or intra-muscular) is much more predictable and has consistently been shown to be greater than an equivalent oral dose (275,276). Thus, in some patients switching to an equivalent dose of subcutaneous methotrexate may improve RA disease control, since the consistent improvement in bioavailability leads to patients receiving an effectively higher dose. The authors of 2 small, retrospective cohort studies have observed loss of disease control when patients were switched to equivalent doses of oral methotrexate following a worldwide shortage of parenteral methotrexate (277,278). Subsequently, 2 randomized and blinded clinical trials have systematically described the relative efficacies of oral and parenteral methotrexate:

- Sixty four patients with active (mean DAS28 = 5.6), established (mean duration = 9.7 years) RA despite oral methotrexate (dose 15-20mg/wk) were switched to intramuscular methotrexate (15mg/wk) (279). After 6 weeks of intramuscular methotrexate, the observed clinical benefits were modest: the mean improvement in DAS28 was 0.42 and only 4 patients had achieved DAS28 <3.2. The remaining 54 patients, who continued to display active disease (DAS28 > 3.2), were randomised to either continue 15mg/wk intramuscular methotrexate or escalate the weekly dose to 45mg/wk. After a further 16 weeks follow-up, the observed clinical improvements remained modest and the response rates were identical for both groups. Within each group, 1 further patient achieved DAS28<3.2, 1 further patient achieved ACR20 response, 5 patients exhibited improvement in DAS28>1.2 and no patients achieved a EULAR Good response. Minor adverse events were reported slightly more frequently in the dose escalation group. Thus, increasing beyond 15mg/wk subcutaneous methotrexate is unlikely to provide additional clinical benefits but may place patients at a slightly higher risk of minor adverse effects.
- Braun et al have demonstrated that patients with RA who are treated with subcutaneous methotrexate achieve significantly better clinical outcomes than those treated with oral methotrexate (280). 384 patients with active (DAS28 > 4) predominantly early (median disease duration 2.1-2.5 months) RA were randomised to receive either 15mg/wk oral or 15mg/wk subcutaneous methotrexate. After 24 weeks follow-up, patients treated with subcutaneous methotrexate exhibited significantly higher ACR20 (sc 78% vs oral 70%; p<0.05) and ACR70 (sc 41% vs oral 33%; p<0.05) responses. A significantly higher proportion of patients with delayed presentations (> 1 year disease duration) of RA who received subcutaneous methotrexate achieved an ACR20 response compared to similar patients treated with oral methotrexate (sc 89 vs oral 63; p<0.05). Furthermore, the study demonstrated that the overall effectiveness of methotrexate could be improved further if patients switched to a subcutaneous preparation and then subsequently increased the dose further. Of 52 patients who failed to achieve an ACR20 response with oral

methotrexate, 16 (30%) subsequently achieved an ACR20 response by switching to subcutaneous methotrexate, and 12 (23%) gained an ACR20 response by increasing the weekly subcutaneous methotrexate dose. Adverse event rates were similar for both groups; though subcutaneous methotrexate was associated with a lower rate of diarrhoea.

Whilst there have been no randomised clinical trials describing the value of parenteral methotrexate in combination with other DMARDs it is feasible to presume that switching to subcutaneous methotrexate, in patients receiving triple DMARD therapy, may increase the overall efficacy of the combination by optimising the therapeutic contribution of the methotrexate component. The study by Lambert et al reported very modest improvements when oral methotrexate was switched to a subcutaneous preparation. However, the initial assessment point was after 6 weeks of subcutaneous therapy, thus may have been too early to detect clinical improvement. Furthermore, in a number of patients the weekly methotrexate dose was actually reduced from 20mg/wk to 15mg/wk. The much larger study by Braun et al has clearly shown that subcutaneous methotrexate produces better clinical outcomes than equivalent oral doses and therefore may be a more effective way of initiating methotrexate therapy. Furthermore, a significant minority of patients who had failed to respond to oral methotrexate experienced improved clinical outcomes by first switching to subcutaneous methotrexate and later optimising their weekly dose.

#### ***Step 4 Combination Therapy with Anti-TNF $\alpha$ Blocking Therapy***

Anti-TNF $\alpha$  blocking therapy can produce profound clinical benefits and arrest erosive progression in RA patients who experience persistently active disease despite non-biologic DMARD therapy. Since a substantial subset of patients will experience adequate responses to non-biologic DMARD therapy, most recent RA treatment guidelines restrict the use of anti-TNF $\alpha$  blocking drugs until after non-biologic DMARD treatment failure is evident (111,125,272). In this way, potentially expensive, and occasionally toxic, therapies are reserved for those patients who have demonstrated the greatest need for aggressive treatment. In reality, there are 5 anti-TNF $\alpha$  blocking agents available (adalumimab (Abbot), certoluzimab (UCB), etanercept (Pfizer), golimumab (Centocor) and infliximab (Schering Plough)); however, since this research has been funded by a research grant from Pfizer UK, those patients who exceed DMARD escalation thresholds despite at least 3 months of subcutaneous methotrexate-DMARD combination therapy will be prescribed etanercept.

Etanercept is a soluble tumour necrosis factor alpha (TNF $\alpha$ ) receptor fusion protein which binds, and inactivates, the pro-inflammatory cytokine TNF $\alpha$ . Randomised, placebo controlled studies have demonstrated that patients with persistently active RA despite DMARD therapy experience significant clinical improvements, and lesser rates of radiographic progression, when etanercept is added to their existing DMARD therapy (160,164). Furthermore, etanercept used in combination with methotrexate has been shown to produce significantly greater clinical responses, and lesser rates of radiographic progression, than using either agent alone (166). Several studies have demonstrated the effectiveness of adding etanercept to non-biologic

DMARD therapy in persistently active, established RA (160,164,166,167), and its specific value in the treatment of active, early RA has been described in a number of other clinical studies:

- Bathon et al randomised 632 patients with recent diagnoses (mean disease duration 11-12 months) of RhF positive, erosive and active RA to receive either etanercept monotherapy (10mg or 25mg twice a week) or oral methotrexate (281). After 12 months follow-up it was evident that the patients who had received 25mg etanercept twice weekly had experienced a more rapid improvement in clinical disease activity measures, significantly greater treatment responses and lesser deterioration in measures of erosive and total radiographic damage compared to patients receiving either methotrexate or 10mg etanercept twice weekly. Patients receiving methotrexate experienced higher rates of all types of infection and laboratory monitoring abnormalities than patients in either etanercept group.
- In the COMET study, 528 patients with active (mean DAS28 = 6.5), early (mean disease duration 9.0 months) RA, who had not received either methotrexate or biologic therapy, were randomised to receive either methotrexate monotherapy or methotrexate-etanercept combination therapy (282). After 52 weeks follow-up, head-to-head comparisons showed that methotrexate-etanercept combination therapy produced more rapid, and statistically greater, improvements in disease activity measures (DAS28). A significantly higher number of etanercept treated patients achieved DAS28 remission (DAS28<2.6: MTX 28% vs MTX-ETAN 50%; p<0.0001), radiographic non-progression (MTX 59% vs MTX-ETAN 80%; p<0.0001) normalisation of functional ability (HAQ DI < 0.5: MTX 39% vs MTX-ETAN 55%; p=0.0004) and were able to remain in their usual employment (MTX 9% vs MTX-ETAN 24%; p=0.004)

The majority of clinical studies of anti-TNF $\alpha$  blocking therapies in early RA have tended to compare the relative efficacies of either commencing newly presenting patients on immediate anti-TNF $\alpha$  blocking therapy or commencing anti-TNF $\alpha$  blocking once a short trial of DMARD monotherapy has proven ineffective. However, the British Society of Rheumatology biologic prescribing guidelines require patients to have tried, and failed, at least two non-biologic DMARDs before they can be considered for anti-TNF $\alpha$  blocking therapy. Hence, in reality, many patients will have tried several different DMARDs, and many will be taking combination therapy, prior to commencing anti-TNF $\alpha$  blocking therapy. Therefore, in order to reflect usual clinical practice, and satisfy the ethos of the BSR prescribing guidelines, this research has chosen to place anti-TNF $\alpha$  blocking therapy as the final stage of the step-up DMARD escalation. Hopefully, this approach will determine how much further leeway there is available to optimise the efficacy of step-up DMARD escalation using traditional, cheaper non-biologic DMARDs and won't delay patients with aggressive RA from receiving timely anti-TNF $\alpha$  blocking therapy if they demonstrate persistently active disease.

Anti-TNF $\alpha$  blocking therapies undoubtedly produce profound, often rapid, improvements in RA and, since many patients have struggled to receive adequate responses with any other agents, treatment is traditionally continued indefinitely. However, continued treatment is also associated

with continued concerns relating to the long-term risks of powerful, immunomodulatory therapy; such as, infection, autoimmune phenomena and cancer. Furthermore, patients with early RA who experience a profound response to anti-TNF $\alpha$  blocking therapy may also experience a fundamental change in the immunopathogenetic processes underlying their RA which may not require them to continue aggressive long-term therapy.

- Quinn et al tested the effectiveness of remission induction with anti-TNF $\alpha$  blocking therapy by randomising 20 patients with early (median symptom duration = 6 – 7.4 months) untreated RA to receive either methotrexate monotherapy or methotrexate-infliximab combination therapy (283). Infliximab was discontinued after 12 months and patients were followed for a further 12 months. As expected, patients treated with infliximab experienced a rapid initial improvement in clinical disease activity measures and significantly higher treatment response rates (ACR70: 70% vs 30%). Importantly, following the discontinuation of infliximab, 70% of patients sustained the initial on-treatment clinical responses and did not require therapy escalation.
- Follow-on subgroup analysis of the BeST study described the progress of patients randomised to Group 4 (initial MTX-infliximab combination therapy) who achieved persistent low disease activity (DAS44<2.4 for at least 6 months) and subsequently discontinued infliximab (284). From the initial 120 patients treated with infliximab, 77 (64%) achieved low disease activity and discontinued treatment. Of these patients, 67 (56% of the whole group; 87% of the discontinuing group) remained in persistent low disease activity and did not need to recommence infliximab. The median duration of infliximab therapy prior to discontinuation was 9.9 months. Of the patient's who experienced a clinical flare, the median interval between treatment cessation and recommencement was 3.7 months

Sheehy et al randomised 24 patients with untreated, early (mean symptom duration 6.3 months) RA to receive either methotrexate monotherapy or methotrexate-etanercept combination therapy (285). Etanercept was discontinued after 24 weeks if patients had achieved clinical remission (DAS28<2.6). Remission rates after 24 weeks treatment were significantly higher in the etanercept group (MTX 35% vs MTX-ETAN 85%). Sixty percent of the patients who had achieved clinical remission with etanercept remained in remission after a further 24 weeks follow-up; whereas, only 30% of the methotrexate treated patients were able to sustain remission

These data suggest that short courses of anti-TNF $\alpha$  blocking therapy may have remission inducing properties, since a significant proportion of patients (56-70%) were able to receive short courses of treatment (6-9 months), achieve the pre-defined disease activity target and then discontinue anti-TNF $\alpha$  blocking therapy without any loss of RA disease control. This approach may optimise the clinical and cost-effectiveness of anti-TNF $\alpha$  blocking therapies by allowing patients to experience the benefits of anti-TNF $\alpha$  blocking therapy whilst avoiding prolonged and costly courses of treatment which might, in some cases, prove hazardous. Thus, since this studies DMARD escalation protocol could lead to participant's qualifying for etanercept at much

lower levels of disease activity than currently sanctioned by any treatment guidelines, patients who do qualify for etanercept will initially receive a time limited treatment course. If remission is achieved and sustained, etanercept will be discontinued after 6 months. If, after discontinuing etanercept, patients subsequently experience a clinically apparent, acute flare of RA etanercept will be restarted indefinitely.

## **2.4 DMARD Escalation Thresholds**

Since the intervention under investigation is the potential efficacy of MSUS to assess global disease activity and guide DMARD escalation decisions, participants are randomly assigned to groups that only differ in the methods used to assess global disease activity. All other aspects of their treatment, and particularly the sequence of DMARD escalation, are identical for both groups.

### **2.4.1 Randomisation Process**

Participants were randomised at the point they consented to participate in the research and remained within the same intervention group for the duration of their participation. The randomisation process aimed to distribute common demographic and disease-associated factors equally between both groups, so as to remove any confounding influence they might otherwise have exerted on the final outcomes

The randomisation process was administered by the Robertson Centre for Biostatistics at the University of Glasgow and was conducted using a telephone-based Interactive Voice Response (IVR) system. Participants were randomly assigned to either a Control (DAS28) or Intervention (MSUS) group; however, the randomisation was adaptive to ensure an equal balance of disease related factors between both groups.

*Control Group* – global disease activity assessment and DMARD escalation threshold based on DAS28 score at time of assessment.

*Intervention Group* – global disease activity assessment and DMARD escalation threshold based upon combination of DAS28 and MSUS findings

Adaptive stratified randomisation (minimisation) techniques - based upon participant's rheumatoid factor status, baseline erosive status and baseline DAS28 – were used to ensure an equal balance (and therefore influence) of these factors between both groups. Essentially, mathematical modelling describes the overall level of imbalance between each group for all these factors; new participants are allocated to the group that is most likely to correct (or reduce) the degree of imbalance. Randomisation was based upon RhF status rather than anti-CCP antibodies since, at the time of design, NHSGGC Immunology Laboratory were unable to guarantee that anti-CCP antibody testing would be available for the duration of the recruitment process

*Blinding:* participants and Dr Dale - who applies the DMARD escalation protocol for both groups - were aware (i.e unblinded) of which intervention group the participant has been assigned to. All other members of the research team, and particularly those responsible for scoring the clinical and radiological outcomes, were blinded to each participant's intervention group and level of DMARD therapy. Hence there is single blinded assessment of research outcomes

## **2.4.2 DMARD Escalation Thresholds for Each Group**

As previously discussed, most current RA treatment guidelines recommend regular disease activity assessment using a composite disease activity score (such as DAS28) aiming for a predefined disease activity level (111,112,124,125,168). For this study, the 'gold standard', to which the additional impact of MSUS assessment will be compared, was considered to be disease activity assessment and DMARD steering using the DAS28. The reasoning underpinning the chosen DMARD escalation thresholds is described in detail in the following sections; furthermore, a diagrammatic representation of the potential decision paths for each group is shown in Figure 8

### ***Control Group***

The numerical value of a patient's DAS28 score at a given time point allows their overall disease activity to be categorised into the following levels (161,165):

High disease activity	DAS28 >5.1
Moderate disease activity	$3.2 \geq \text{DAS28} \leq 5.1$
Low disease activity	$2.6 \leq \text{DAS28} < 3.2$
Clinical remission	DAS28 < 2.6

The thresholds for DAS28 and DAS44 groupings were developed from the consensus opinion of groups of rheumatologists who were asked to decide whether certain patient profiles qualified for DMARD escalation (180). High disease activity was defined as clinical profiles that, in the opinion of the participating rheumatologists, qualified for DMARD escalation. Low disease activity was defined as clinical profiles that would allow DMARD tapering and/or cessation. The discriminatory ability of individual variables, and groups of variables, to differentiate between high and low disease activity states was then determined and these weightings form the basis of the current DAS28 and DAS44 calculations (180,286). Whilst DAS28 and DAS44 scores provide useful, easily understood numerical measures of global disease activity a number of important points relating to their derivation must be considered:

1. the initial definition of high and low disease activity was based on the clinical judgements of a group of rheumatologists. Clinical practice and interpretation of clinical findings varies significantly between individual clinicians; thus, the assumptions underpinning current

DAS28 and DAS44 disease activity definitions – and all subsequent treatment response criteria - were inherently subjective, and at risk of inter-observer variability, even before specific values had been derived

2. a specific definition of moderate disease activity was not described. Moderate disease activity is inferred in patients who achieve DAS28 and DAS44 scores that fall between the numerical thresholds for either high or low disease activity. By extension this should refer to those patients who, in the opinion of the assessing rheumatologists, neither required DMARD escalation nor reduction. However, in practice this assumed definition is not entirely accurate since patients with moderate disease activity are usually considered to have some degree of active inflammatory disease and therefore do require DMARD escalation

For the purposes of this study, the control group DMARD escalation threshold will be a DAS28 greater than, or equal to, 3.2 (i.e at least moderate disease activity) at least three months after DMARD escalation. This threshold is in keeping with current practice in existing NHSGGC early arthritis clinics and is equivalent to similar thresholds employed by several preceding early RA step-up DMARD strategy studies (118,145,146,252).

### ***Musculoskeletal Ultrasound Group***

Currently, there remains no consensus regarding which joints need to be examined by MSUS to adequately assess global disease activity (i.e the minimal joint set); nor is there an agreed MSUS definition of active synovitis. The following section will describe the development of this studies MSUS joint set and thresholds for DMARD escalation:

#### **Definition of MSUS Pathology**

The Outcome Measures in Rheumatology (OMERACT) and EULAR MSUS Working Parties have published consensus definitions to improve accuracy and description of MSUS findings (287). These definitions are widely used and were adopted by this study:

*Synovial Fluid* – abnormal hypoechoic or anechoic intra-articular material that is displaceable and compressible; does not exhibit Doppler signal

*Synovial Hypertrophy* – abnormal hypoechoic intra-articular tissue that is non-displaceable and poorly compressible; may exhibit Doppler signal

*Tenosynovitis* – hypoechoic or anechoic thickened tissue, with or without fluid within the tendon sheath, that is seen in two perpendicular planes; may exhibit Doppler signal

*Enthesopathy* – abnormally hypoechoic (loss of normal fibrillar architecture) and/or thickened tendon, or ligament, at its bony attachment seen in two perpendicular planes; may exhibit Doppler signal and/or bony changes, including enthesophytes, erosions or irregularity

*Bone Erosion* – an intra-articular discontinuity of the bone surface that is visible in two perpendicular planes

Since comparative studies of MSUS findings, synovial histology and clinical outcomes have demonstrated that the presence of intra-articular Doppler signal is strongly associated with active inflammatory disease (whether by demonstration of an active inflammatory infiltrate, radiographic evidence of erosive progression or increased frequency of acute flare – see Section 1.5) active synovitis in a single joint requires demonstration of intra-articular Doppler signal. Equally, since the primary articular lesion of RA is synovitis, and since this study is primarily focussed on assessing measures of global inflammatory disease activity (and not their erosive sequelae), MSUS assessment will focus on findings of synovial hypertrophy (and/or effusion), and intra-articular Doppler signal and exclude tenosynovitis, enthesopathy and bone erosions. In this way, the MSUS assessment method will be deliberately focused on identifying evidence of active inflammatory disease, quicker to apply and potentially easier to apply during normal clinical practice

### **Grading of MSUS Findings**

Without the use of time-consuming digital image analysis software grading of MSUS findings during routine clinical practice remains largely subjective; being dependent upon the examiners ability to both acquire and interpret the individual components of each MSUS image. Several semi-quantitative grading systems have been proposed whereby standard definitions of the extent of MSUS findings attach a numerical value to the observed findings (288,289). However, there is not yet a universal consensus on which is the correct grading system, nor has their ability to influence outcomes in longitudinal clinical research been described. At the time that this protocol was being devised, the grading system proposed by Szkudlarek et al was most widely used (288). This grading system was originally based on MSUS findings in MCP and MTP joints; though for the purposes of this research the grading principles were extended to the radiocarpal and PIP joints as well (see below).

*Synovial Hypertrophy* – non-compressible hypoechoic intra-articular area (synovial thickening – see Figure 5)

Grade 0	no synovial thickening
Grade 1	minimal synovial thickening; filling the angle between the periarticular bones, without bulging over the line linking tops of the bones
Grade 2	synovial thickening bulging over the line linking tops of the periarticular bones but without extension along the bone diaphysis

Grade 3      synovial thickening bulging over the line linking tops of the periarticular bones and with extension to at least one of the bone diaphysis

*Joint Effusion* – compressible anechoic intra-articular area (see Figure 6)

Grade 0      no effusion

Grade 1      minimal amount of joint effusion

Grade 2      moderate amount of joint effusion (without distension of the joint capsule)

Grade 3      extensive amount of joint effusion (with distension of the joint capsule)

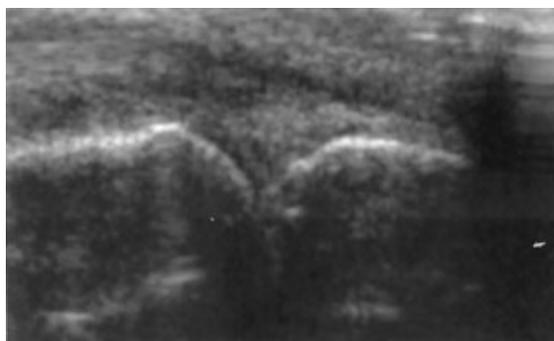
*Doppler Signal* – extend of Power Doppler signal identified within the synovium (see Figure 7)

Grade 0      no flow in the synovium

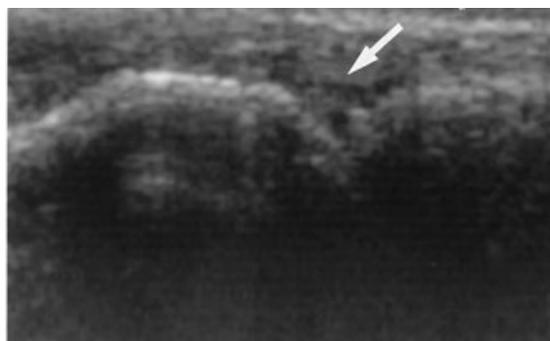
Grade 1      single vessel signals

Grade 2      confluent vessel signals in less than half the area of the synovium

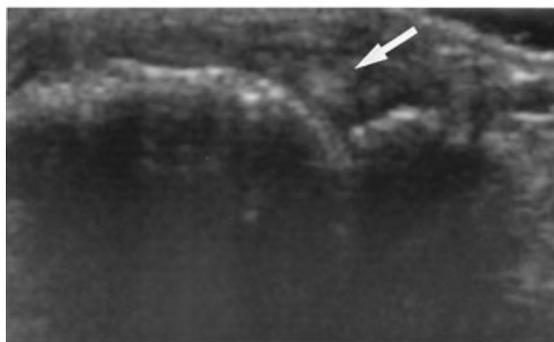
Grade 3      vessel signals in more than half the area of the synovium



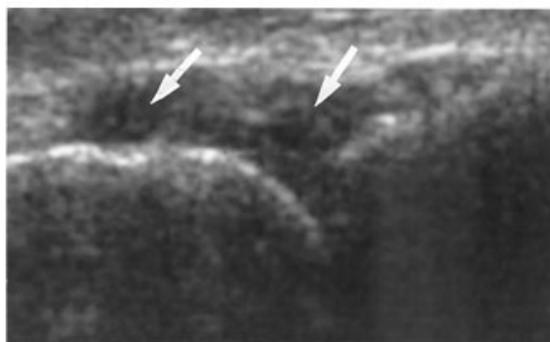
Grade 0



Grade 1



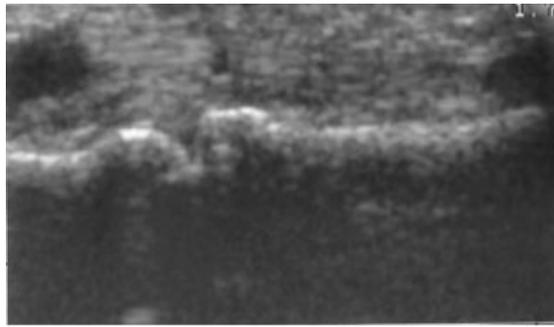
Grade 2



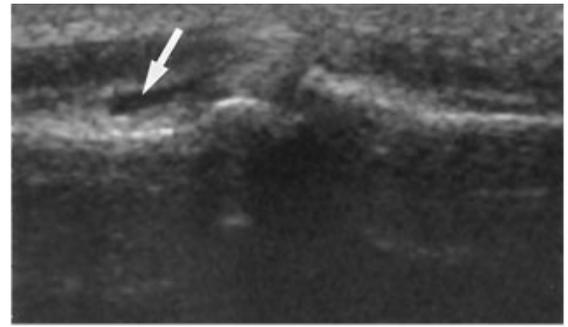
Grade 3

**Figure 5:** Grading of Synovial Hypertrophy  
From: Szkudlarek et al. *Arthritis and Rheumatism* 2003; 48: 955-962

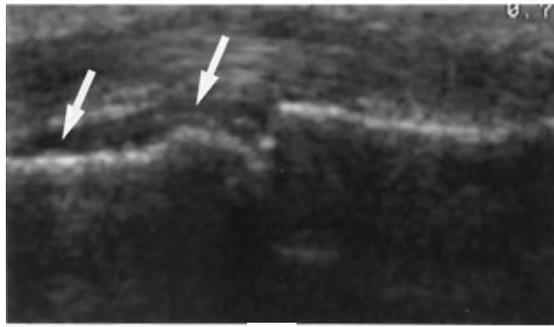
---



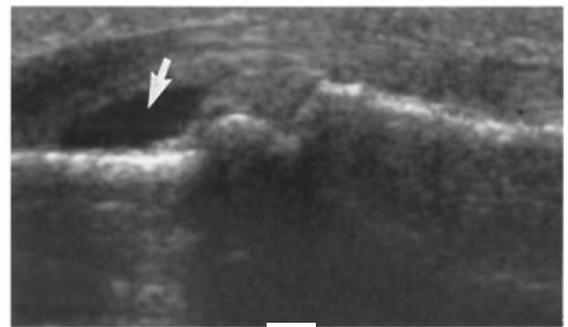
Grade 0



Grade 1



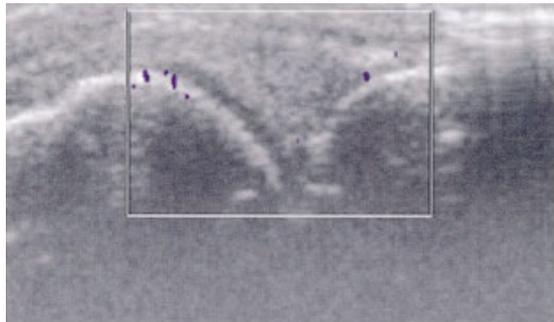
Grade 2



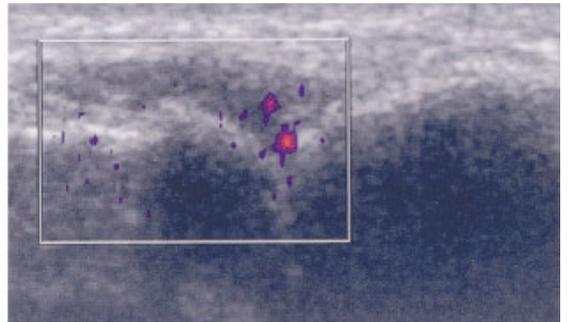
Grade 3

**Figure 6:** Grading of Joint Effusion

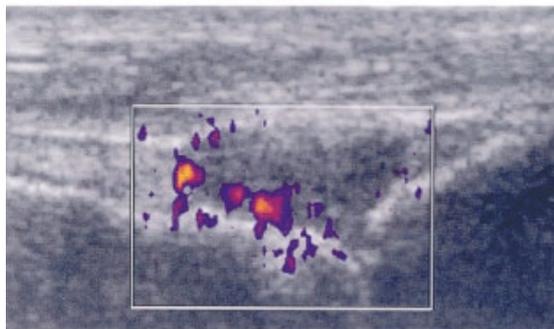
From: Szkudlarek et al. *Arthritis and Rheumatism* 2003; 48: 955-962



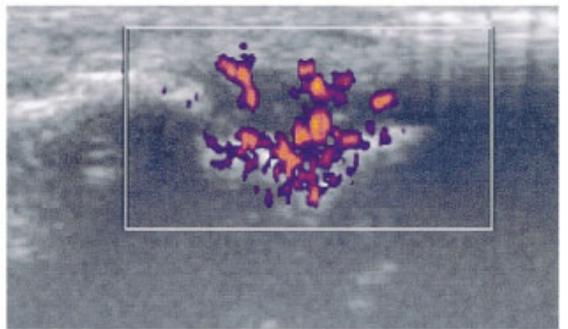
Grade 0



Grade 1



Grade 2



Grade 3

**Figure 7:** Grading of Power Doppler Signal

From: Szkudlarek et al. *Arthritis and Rheumatism* 2003; 48: 955-962

### **Development of Minimal MSUS Joint Set**

In order for MSUS to be directly applicable to clinical practice, the methods used need to be both reliable and efficient. Whilst scanning large numbers of joints in detail will produce findings which most accurately reflect global RA disease activity, it is also likely to be time consuming and impractical in a busy clinical setting. Thus it is important to identify and test the minimum number of joints requiring MSUS examination to balance both accuracy and practicalities. Several clinical studies have described how different permutations of minimised joint sets correlate to larger, more comprehensive joint sets and their various findings will form the basis of the joint set used throughout this research:

- Scheel et al systematically examined the clinically more affected 2<sup>nd</sup> to 5<sup>th</sup> MCP and PIP joints of 46 patients with RA for evidence of synovial hypertrophy and joint effusion (290). The presence of Doppler signal was not reported. Synovial hypertrophy and joint effusion were graded as a single finding using a semi-quantitative system (grade 0-3) adapted from that proposed by Szkudlarek et al (288). Total scores for different combinations of joints were compared using receiver operating characteristic (ROC) and area under the curve (AUC) analyses. Scores generated using the findings at 2<sup>nd</sup> through 5<sup>th</sup> MCP joints produced the lowest AUC (0.69), sensitivity (45.7%) and specificity (90%). There was no statistically significant difference between the scores generated using all other joint combinations since each generated high AUC values: 2<sup>nd</sup> through 5<sup>th</sup> MCP and PIP joints (AUC 0.90, Sensitivity 0.76, Specificity 0.90); 2<sup>nd</sup> through 4<sup>th</sup> MCP and PIP joints (AUC 0.90, Sensitivity 0.80, Specificity 0.90); 2<sup>nd</sup> and 3<sup>rd</sup> MCP and PIP joints (AUC 0.85, Sensitivity 0.61, Specificity 1.0) 2<sup>nd</sup> through 5<sup>th</sup> PIP joints (AUC 0.90, Sensitivity 0.80, Specificity 1.0).
- Naredo et al performed comprehensive MSUS examinations on 160 patients with active, established RA prior to, and after 6 months of, biologic therapy (226). Systematic examination of grey scale and Power Doppler signal findings was performed in 44 peripheral large and small joints and all findings were graded semi-quantitatively. Different combinations of reduced joint sets were produced based upon the frequency of finding synovial and Power Doppler abnormalities at baseline. A reduced 12 joint model was then produced using only those joints which detected at least 90% involvement by synovitis and Power Doppler signal. The reduced 12 joint model comprised examining the elbow, wrist, 2<sup>nd</sup> and 3<sup>rd</sup> MCP, knee and ankle bilaterally for evidence of grey scale synovitis and Power Doppler signal. The 12 joint set identified 100% of patients with grey scale synovitis on the full 44 joint set and 94.4% of patients with Power Doppler signal. Furthermore, a simplification which focussed solely on Power Doppler findings in the 12 joint set demonstrated strongly positive, and statistically significant correlations with the corresponding total Power Doppler scores and indexes found in the full 44 joint set (Pearson Correlation Coefficients: Joint Count = 0.89,  $p < 0.0005$ ; Joint Index = 0.90,  $p < 0.0005$ ).
- Backhaus et al have proposed a 7 joint set of the clinically dominant hand and foot (291). This joint set comprises the dominant wrist, 2<sup>nd</sup> and 3<sup>rd</sup> MCP, 2<sup>nd</sup> and 3<sup>rd</sup> PIP and 2<sup>nd</sup> and 5<sup>th</sup> MTP joints. The derivation of the joint set is not explicitly described in the original publication.

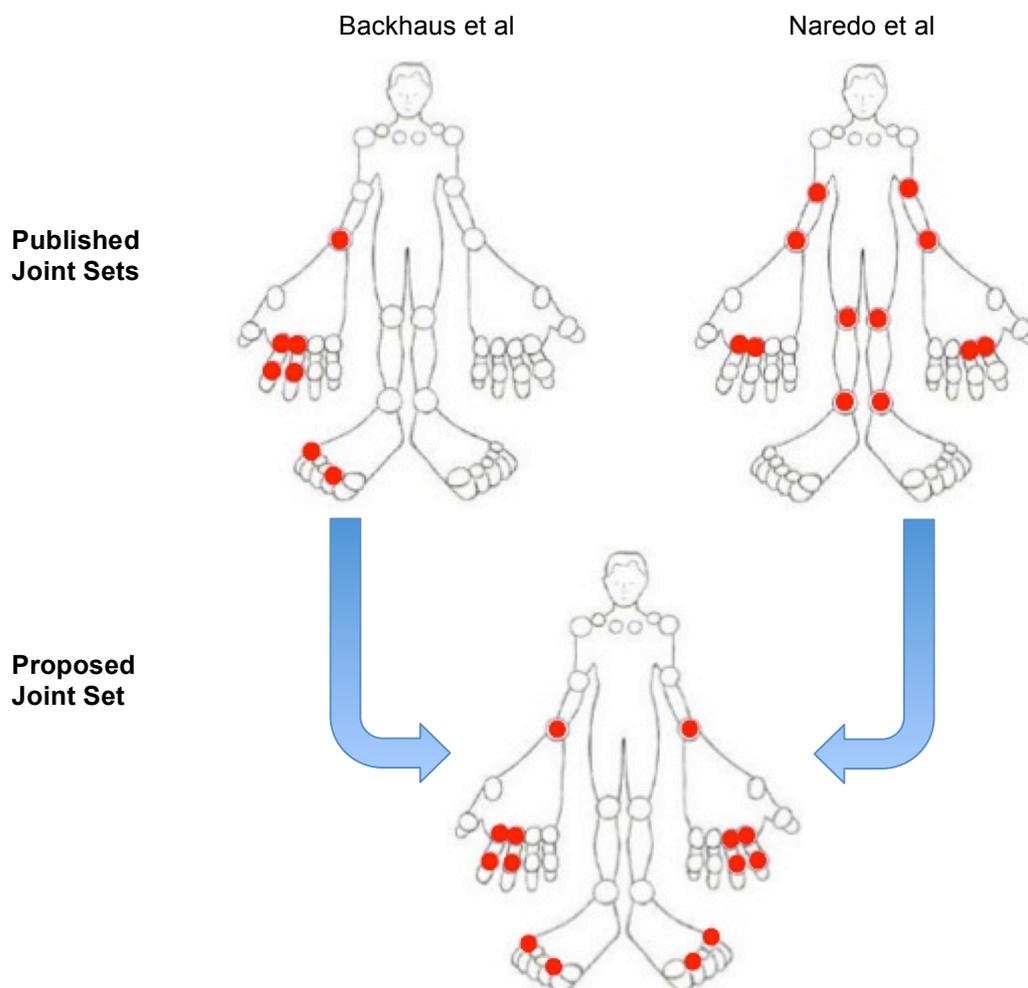
Nevertheless, a validation exercise did demonstrate that the score correlates positively with clinical disease activity measures and is responsive to changes in therapy. One hundred and twenty patients with RA or psoriatic arthritis underwent MSUS examination of the 7 joint set immediately prior to, 3 months after and 6 months after a significant change in DMARD or biologic therapy. Grey scale and Power Doppler findings for synovitis, tenosynovitis and erosions were graded semiquantitatively (0-3). Mean baseline Power Doppler synovitis score was 3.3, suggesting a relatively low overall number of joints with active synovitis. Mean grey scale and Power Doppler scores for synovitis fell significantly 6 months after a therapy change (Grey scale synovitis: Baseline 8.1 to 5.5,  $p < 0.05$ ; Power Doppler synovitis: Baseline 3.3 to 2.0,  $p < 0.05$ ). There were statistically significant, moderately positive correlations observed between the change in mean Grey scale and Power Doppler synovitis scores and corresponding changes in DAS28 after 6 months of treatment (Grey scale synovitis vs DAS28:  $r = 0.38$ ,  $p < 0.05$ ; Power Doppler synovitis vs DAS28:  $r = 0.31$ ,  $p < 0.05$ ). Furthermore, reliability analyses demonstrated moderate inter-reader agreement for both semiquantitative assessment of grey scale synovitis ( $\kappa = 0.55$ ) and Power Doppler synovitis ( $\kappa = 0.67$ ) and moderate intra-reader agreement for semiquantitative assessments ( $\kappa = 0.64$ ).

A recent longitudinal clinical study has demonstrated that the 7 joint set is sensitive to changes in clinical and laboratory measures of RA disease activity in a large cohort of RA patients (292). Four hundred and thirty patients underwent clinical (DAS28), laboratory (CRP and ESR) and MSUS 7 joint set assessment of global disease activity prior to, 3, 6 and 12 months after changing their DMARD/biologic treatment. Twelve months after treatment changes, mean grey scale and Power Doppler synovitis scores had fallen significantly for all groups. Patients receiving first or second line DMARDs demonstrated a significant correlation between the change in grey scale and Power Doppler synovitis scores and the corresponding change in DAS28 (1<sup>st</sup> Line DMARD: GS  $r = 0.419$ ,  $p < 0.001$ ; PD  $r = 0.459$ ,  $p < 0.001$ . 2<sup>nd</sup> Line DMARD: GS  $r = 0.257$ ,  $p = 0.008$ , PD  $r = 0.283$ ,  $p = 0.007$ ). A significant, positive correlation between changes in MSUS findings and ESR and CRP levels was demonstrated for patients receiving first line biologic therapy (ESR: GS  $r = 0.207$ ,  $p = 0.011$ ; PD  $r = 0.179$ ,  $p = 0.032$ . CRP: GS not quoted; PD  $r = 0.312$ ,  $p < 0.001$ ), though results for DAS28 were not reported. Overall, the magnitude of change in MSUS scores appeared greater than the corresponding change in DAS28.

- Dougadas et al have reported the clinimetric properties and reliability of several MSUS joint sets to quantify global RA disease activity in 76 patients with active disease requiring biologic therapy (293). Overall, the joint sets chosen by Dougadas et al comprised substantially more joints (20-38) than those proposed by Scheel et al, Backhaus et al and Naredo et al. Each of the chosen joint sets (20 joints: 1<sup>st</sup>-5<sup>th</sup> MCP and MTP bilaterally. 28 joints: DAS28 joint set. 38 joints: DAS28 joint set + 1<sup>st</sup>-5<sup>th</sup> MTPj bilaterally) demonstrated good intra-observer agreement (intra-class correlation coefficient range: Grey scale 0.85 – 0.97, Power Doppler 0.61 – 0.96), construct validity (Alpha Cronbach test: Grey scale 0.76 – 0.89, Power Doppler 0.81 – 0.86) and sensitivity to change after commencing TNF-alpha blocking therapy. These findings were comparable, if not better, than clinical assessments performed in similar joint sets though

unfortunately correlation analyses between the different examination methods and joint sets were not reported.

- Filer et al performed a longitudinal clinical study of 58 patients with undifferentiated inflammatory arthritis to describe the ability of different MSUS joint set combinations to predict a future diagnosis of RA (146,151). Whilst this is a study of the diagnostic ability of MSUS joint sets, and there is no comparison to alternative disease activity measures, a number of important factors were illustrated. When used in conjunction with the Leiden predictive criteria for RA, MSUS grey scale and Power Doppler findings significantly improved the examiners ability to diagnose RA. Specifically, logistic regression analyses demonstrated that grey scale and Power Doppler involvement of the wrist, MCP and MTP joints and symmetrical involvement of the wrist and MTP joints provided additional, independent predictors of RA over and above the Leiden score. Moreover, combining MSUS findings into a 10 joint set, based upon the 12 joint set proposed by Naredo et al (see above) but with the knee joints excluded, further increased the AUC and sensitivity of the assessment.



**Figure 8:** Derivation of limited joint set

The previously described studies highlight several important factors that were taken into account when the joint set used by this study was being planned:

1. Minimised MSUS joint sets, often restricted to easily accessible peripheral small joints, are comparable to larger, global MSUS joint sets
2. Minimised MSUS joint sets are sensitive to change following commencement, or escalation of DMARD and/or biologic therapy
3. Simplified systems which focus on the presence of Power Doppler signal provide comparable findings to those which focus on Power Doppler signal and grey scale findings
4. Bilateral involvement of wrist and MTP joints is predictive of RA

Since there were no universally accepted minimal joint sets available, a joint set pragmatically combining the important properties of those that had already been proposed (Scheel et al, Backhaus et al, Naredo et al) and had undergone a degree of validation was developed (Figure 8). RA classically presents with symmetrical joint involvement, and since Naredo et al (226), and later Filer et al (151), emphasised the importance of assessing bilateral joint involvement, this study also deliberately chose to use a symmetrical joint set. The joints chosen are the bilateral extrapolation of the unilateral set proposed by Scheel et al and Backhaus et al (290-292) and include several joints in common with the set proposed by Naredo et al. The joint set proposed by Naredo et al also included several large peripheral joints (elbow, knees, ankles). These were excluded since scanning of these regions is more complicated, potentially time consuming and requires a higher degree of MSUS expertise that might limit the adoption of the proposed joint set by less experienced rheumatology sonographers. Prior to commencement of the research period the components of the proposed 14 joint MSUS set, and MSUS definitions of active RA, were discussed with Professor Phillip Conaghan (Professor of Musculoskeletal Medicine, University of Leeds), a rheumatologist with a noted interest and reputation in the use of modern imaging techniques to assess disease activity in rheumatological conditions. Indeed, the proposed 14-joint count was recently tested alongside an alternative, similarly structured, 18-joint count currently being validated by the OMERACT-Ultrasound Task Force (the Global Synovitis Score) and was found to have similar metric properties (294)

### **MSUS Disease Activity Assessment**

Whenever MSUS disease activity assessment was indicated the following joints were examined systematically: bilateral index and middle PIPj, index and middle MCPj, wrist, 2<sup>nd</sup> and 5<sup>th</sup> MTPj. Each joint was examined in the dorsal, longitudinal plane with the participant's hands and feet resting in the neutral position. The presence of synovial hypertrophy and PD signal was graded using the semi-quantitative system proposed by Szkudlareck et al (288). The nature of unclear or equivocal findings was confirmed by examination in the transverse plane. If there was no PD signal evident during dorsal examination of the MCP and PIP joints the palmar aspect of the joint was also examined. Appendix D includes an example of how MSUS findings were recorded during each assessment

## Ultrasound Machine Settings

Interpretation of images gathered by MSUS is subject to a degree of subjective, inter-observer variability which can, in some cases, lead to erroneous interpretation of the final image.

Furthermore, individual ultrasound machines differ in how they present the same image and the content of the final image can be influenced by variations in a machine's standard settings.

Therefore, the following steps were taken to try and minimise any additional inherent variability associated with the MSUS assessment and therefore restrict any variability this might subsequently introduce in DMARD escalation decision.

1. A single, portable machine, and single linear probe with standardised settings, was used for all MSUS assessments.
2. A single researcher (Dr James Dale) performed all MSUS assessments using a standardised technique
3. For each assessment, predefined, standardised ultrasound machine settings were used.

A portable Voluson I (GE Healthcare, UK) with a 10-16mHz linear probe (SP 10-16RS, GE Healthcare, UK) was purchased solely for use during this research. This machine and probe were used for every single MSUS assessment. For each assessment predefined, standardised ultrasound machine settings were used. In practice, these were identical to the pre-programmed settings provided by the manufacturer to optimise MCPj assessment.

### *Gray Scale Settings*

Frequency	16.0-13.5 (high resolution setting)
Focus point	Single point – placed in line, or just below, region of interest
Depth	2.2cm
Power	100%
Zoom	100%
Gain	Variable – adjusted to examiner's preference to obtain subjective highest quality image
Persistence	1
Dynamic Contrast	4
Edge Enhance	3
Line Filter	2
Quality	High
Speckle Reduction Imaging (SRI II)	High

### *Power Doppler Settings*

Frequency	High
Pulse Repetition Frequency	0.9kHz
Wall Filter	Low
Region of interest box – size and shape adjusted to incorporate the whole of the intra-articular space, surrounding articular surfaces and shafts of bone and the upper margin of image	

Gain – variable and adjusted to examiner's preference to obtain highest quality images with least evidence of artefact. PD gain was set at the level just below that which caused abnormal Doppler artefact to appear below the surface of bone

## **Musculoskeletal Ultrasound Definition of Active Disease and Thresholds for DMARD**

### **Escalation**

#### *Timing of MSUS Assessment in Relation to Clinical Assessment*

The positioning of MSUS assessment in relation to DAS28 assessment, and when it's most likely to positively influence DMARD escalation decisions, needs to be carefully considered. There will be some scenarios where the additional information provided by MSUS assessment is unlikely to influence treatment decisions since the decision can be reached on clinical grounds alone.

Furthermore, there will be some scenarios where the additional information provided by MSUS will lead to a different treatment decision than that suggested by the clinical findings..

The following clinical scenarios were used to consider how best to place MSUS assessment alongside DAS28 assessment:

- DAS28 > 5.1 – corresponds to high disease activity. This level of disease activity is least likely to have been influenced by external factors and is most likely to represent active inflammatory disease. DMARD escalation is indicated based upon the clinical findings and therefore MSUS assessment is not indicated
- $3.2 \leq \text{DAS28} < 3.2$  and at least 2 clinically swollen joints – a moderate disease activity score and demonstration of at least 2 clinically swollen joints provides strong clinical evidence of persistent inflammatory disease activity. Therefore, DMARD escalation is indicated based on clinical findings and MSUS is not required since it is unlikely to influence the clinical decision
- $3.2 \leq \text{DAS28} < 5.1$  and one or no clinically swollen joints – in this scenario the DAS28 score is elevated above the traditional DMARD escalation threshold, however, there is minimal clinical evidence of active synovitis. MSUS assessment is indicated to determine whether the clinical findings represent active subclinical synovitis and therefore require DMARD escalation. If subclinical synovitis is not identified it is highly probable that the clinical findings (and any on-going joint symptoms) are not related to active synovitis and would not respond to further DMARD escalation. Exclusion of subclinical synovitis should prompt an alternative treatment approach in any patients who remain symptomatic. Furthermore, symptomatic patients without active synovitis will avoid exposure to the potential risks of unnecessary additional DMARD therapy
- DAS28 < 3.2 – corresponds to low disease activity and in many situations is the disease activity target at which DMARD therapy is aimed. MSUS assessment is indicated to identify persistent subclinical synovitis. Confirmation of subclinical synovitis should prompt further DMARD escalation, even in asymptomatic patients. Thus, MSUS findings have supported further intensification of DMARD therapy, over-and-above that suggested by the DAS28 assessment.

### *MSUS Definition of Active RA*

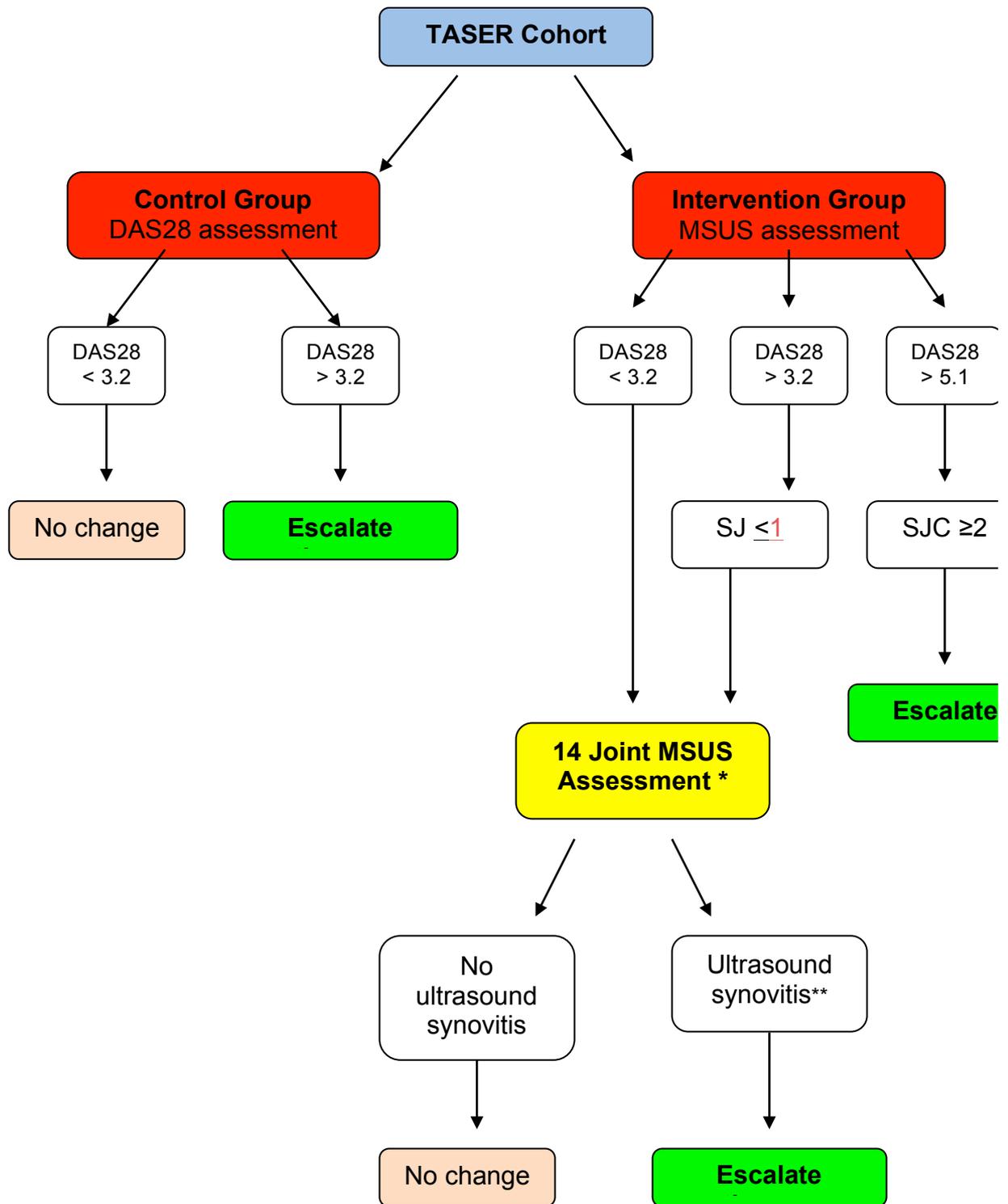
Whilst RA is a systemic disease, MSUS assessment initially generates findings that represent the individual joints being examined rather than global disease activity. Findings for individual joints need to be collated into a single measure of global disease activity that can then be used to inform DMARD escalation decisions. Since there is not yet a universally accepted reduced joint set there is not (yet) a universally accepted MSUS definition of active RA. In the development of this research a number of presumptions were made which informed the development of a unique MSUS definition of active RA based on evidence discussed in the preceding sections:

1. The proposed limited joint set will identify a greater burden of active synovitis than that suggested by clinical assessment
2. The proposed limited joint set will exclude on-going synovitis in a subset of patients whose DAS28 assessment is biased by non-inflammatory joint disease
3. The presence of intra-articular Power Doppler is abnormal and corresponds to the presence of active synovitis. Conversely, regardless of grade, grey scale synovial hypertrophy without Power Doppler signal represents chronic synovial hypertrophy rather than active synovitis
4. An increase in global disease activity will be represented by abnormal MSUS findings, and specifically abnormal Power Doppler findings, in at least 2 or more of the joints comprising the reduced joint set
5. Erosions represent the sequelae of preceding active synovitis; therefore, identification of synovial hypertrophy and intra-articular Power Doppler signal will provide a more accurate representation of global disease activity at a given time point

Thus, the following definition of active RA, based upon MSUS findings from the proposed limited joint set, was adopted:

“The presence of at least grade 1 or higher intra-articular Power Doppler signal in at least 2 joints examined by MSUS for grey scale synovial hypertrophy and Power Doppler signal”

This definition was chosen because: 1. it was in keeping with recent evidence that RA patients in low disease activity who still exhibited 2 clinically swollen joints were at an increased risk of future erosive progression compared to patients with no swollen joints (192) 2. it suggests active synovitis is not confined to a single joint and 3. reduced the risk of PD artefact within a single joint being misinterpreted as active synovitis and contributing to erroneous DMARD escalation (224,295)



\* MSUS Assessment – grey scale and power Doppler assessment of index and middle PIPj hands bilaterally, index and middle MCPj hands bilaterally, radio-carpal joint bilaterally and 2<sup>nd</sup> and 5<sup>th</sup> MTPj feet bilaterally

\*\* Ultrasound Synovitis – power Doppler signal  $\geq 1$  affecting 2 or more joints

**Figure 9:** Diagrammatic representation of potential DMARD escalation decision pathways for Control (DAS28) and Intervention (MSUS) groups; based upon possible clinical and MSUS scenarios

## 2.5 Outcome Measures

A broad range of clinical, laboratory and radiological outcome measures were collected at regular intervals throughout the follow-up period. Since only members of the MSUS group underwent MSUS intervention it was not possible to blind the participants to their randomisation group. However, Sister Anne Stirling, who collected all the clinical outcomes, and the 2 radiologists who will grade the radiological outcomes, were kept fully ignorant of each participant's randomisation group so that the assessment and grading of the main clinical and radiological outcomes is considered to be 'single blinded'. To avoid treatment bias, Dr Dale was not made aware of any of the outcome assessments until the study had fully completed. Furthermore, none of the outcome measures were used to influence a participant's treatment decision. Table 4 summarises which outcomes were collected at each time point throughout the research.

	Assessor	Blinded	Month																					
			Screening	Baseline	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18		
RF/CCP	NHSGGC	Y	X																					
DAS28	JD	N		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
Ultrasound*							X			X			X			X			X					
DAS44	AS	Y	X	X		X			X			X			X			X				X		
Pain VAS				X		X			X			X			X			X					X	
Physician global VAS				X		X			X			X			X			X						X
HAQ score				X		X			X			X			X			X			X			X
EQ5-D				X		X			X			X			X			X			X			X
Plain X-ray hands/feet	Rad x 2	Y	X																				X	
MRI hand/wrist				X																				X
Biomarkers	Lab	Y		X			X									X								
FBC/ESR	NHSGGC	Y	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
U&E/LFT			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CRP			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Adverse events	JD	N			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	

**Table 4:** Timing of collection of each outcome measure (NHSGGC = NHSGGC Laboratory Services, JD = Dr James Dale, AS = Sr Anne Stirling, RADx2 = independent radiologists, Lab = collaborating scientists)

### 2.5.1 Clinical Outcome Measures

Sister Anne Stirling, rheumatology research nurse and metrologist, accompanied Dr Dale to research clinics at each of the participating hospitals and collected all clinical outcome data.

Sister Stirling assessed participants after Dr Dale in a clinic room sufficiently far from Dr Dale's to prevent her overhearing any of the preceding treatment discussions. To ensure Sister Stirling remained blinded to assessment groups, participants were clearly instructed to avoid any discussions of their RA status, clinic consultation and whether or not they underwent ultrasound assessment. The ultrasound machine was present at every single consultation and Dr Dale attempted to ensure that all clinic visits were of approximately equal duration. The following assessments of clinical status and functional ability were collected at baseline and every three months thereafter until study completion

**44 Joint Disease Activity Score (DAS44)** (286,296): The DAS44 is a composite measure of disease activity derived by imputing the values of the 44 swollen joint count (44SJC), Ritchie Articular Index (RAI), patient global health 10cm VAS and ESR into the following equation:

$$\text{DAS44} = (0.53938 \times \sqrt{\text{RAI}}) + (0.6465 \times 44\text{SJC}) + (0.33 \times \ln \text{ESR}) + (0.00722 \times \text{GlobalVAS})$$

DAS44 values at a single time point allows global disease activity to be categorised into various levels of severity:

Remission	DAS44 < 1.6
Low disease activity	DAS44 ≤ 2.4
Moderate disease activity	2.4 < DAS44 ≤ 3.7
High disease activity	DAS44 > 3.7

Furthermore, EULAR response criteria use the relationship between the net change in DAS44 and the final DAS44 value to categorise qualitatively a patient's apparent response to therapy (297):

DAS44 at endpoint	Improvement in DAS44		
	≤ 1.2	> 0.6 and ≤ 1.2	≤ 0.6
≤ 2.4	Good	Moderate	None
>2.4 and ≤ 3.7	Moderate	Moderate	None
>3.7	Moderate	None	None

**Table 5:** Derivation of EULAR response using DAS44 thresholds

**Health Assessment Questionnaire Disability Index (HAQ):** The HAQ is widely considered as the gold standard measure of functional ability. It was originally developed in patients with arthritis (298) but since has been validated as a measure of functional ability in patients with a wide range of arthritic and chronic conditions (299). The HAQ is administered as a self, or assessor, completed questionnaire where statements relating to activity of daily living are graded using semi-quantitative Likert Scales (0 – without difficulty, 1 – with some difficulty, 2 – with much difficulty, 3 – unable to do). Scores are adjusted for requiring external assistance and the use of physical aids. The final score is between 0-3.0 with increasing scores representing worse functional ability. The minimally clinically important change is approximately 0.22 and functionally independent patients can report mildly elevated HAQs between 0.38 and 0.45 (300). Appendix E displays the HAQ questionnaire proforma.

**EQ-5D-3L Questionnaire (Euro-QOL):** The EQ-5D-3L questionnaire and EQ VAS provide a generic, highly simplified measure of health status (301). Domains relating to mobility, self care, usual activities, pain/discomfort and anxiety/depression are graded using a 3 level cardinal scale (1 – no problems, 2 – some problems, 3 – extreme problems). A 20cm vertical VAS provides a numerical depiction of the respondent’s general health (0 – the best health you can imagine; 100 – the worst health you can imagine). For individual patients, EQ-5D-3L results can be represented as a health profile by listing the responses for each health domain with the numerical value of the global VAS. Alternatively, a numerical health index can be calculated by using predefined value sets to adjust for the weighting of each domain and global VAS. On a group level, the EQ-5D-3L can be summarised by the frequency of patients responding to the different levels of each domain or as a description of the range and central tendency of the global VAS and EQ-5D-3L health index

**Pain 10cm Visual Analogue Score:** Ten centimetre pain visual analogue scores provide an easily understood, and commonly used, measure of a patient’s symptom burden at a single time point. Its change over time provides a numerical representation of the change in a patient’s symptoms in response to therapy. However, pain VAS are non-specific and are often influenced by other, non-RA related, causes of pain

Taken together this group of clinical variables comprises the minimum, core-set recommended by the American College of Rheumatology to standardise descriptions of outcomes in longitudinal clinical studies (302)

## **2.5.2 Laboratory Outcome Measures**

**Acute Phase Measures:** CRP and ESR were measured at each monthly visit and were analysed by the routine methods of NHSGGC Laboratory services. For reporting purposes, values corresponding to each of the 3 monthly clinical assessments will be reported

**Monitoring Blood Tests:** To comply with national DMARD monitoring guidelines, samples for testing of U+E, LFT and FBC were analysed each month by NHSGGC Laboratory services. Whilst these blood tests can sometimes demonstrate abnormalities related to RA activity their values will not be reported as an outcome measure.

## **2.5.3 Radiological Outcome Measures**

All images for radiological outcome analysis were collected at baseline and after 18 months using local NHSGGC radiology facilities and each department’s standard image acquisition methods. The degree of change in standardised radiological scores of joint damage will serve to describe the rate of erosive progression for each patient over the follow-up period of the study. There will be two independent radiological outcome measures collected:

**MRI Dominant Hand and Wrist:** This will serve as the primary radiological outcome measure and images will be graded using the OMERACT Rheumatoid Arthritis MRI Scoring system

(RAMRIS) (250,303-305). MRI has several advantages over traditional plain radiographs: 1. MRI is more sensitive than plain radiographs at detecting erosions, 2. MRI allows direct visualisation of synovitis (and other peri-articular pathology) and 3. MRI evidence of bone marrow oedema has been consistently shown to predict future erosive joint damage (306). To reduce potential inter-machine bias all MRI scans were performed using the same MRI machine (1.5T Siemens Avanto, Glasgow Royal Infirmary Department of Radiology) and scanning protocols that had been standardised by the local radiographers. The dominant wrist and 2<sup>nd</sup>-to-5<sup>th</sup> MCPj were imaged using the following sequences: T1-weighted - before and after intravenous gadolinium contrast – and T2-weighted fat saturated. In order to limit examination time, scanning was restricted to the dominant wrist since studies comparing unilateral and bilateral combinations of joints have not demonstrated any significant difference between either combination's ability to detect structural progression (307)

*The OMERACT RAMRIS System:* The RAMRIS system was developed by consensus to standardise acquisition and grading of MRI images for research purposes and to facilitate its use as an outcome measure (303). The RAMRIS system semi quantitatively grades the extent of MRI erosions, synovitis and bone marrow oedema affecting each joint region individually. Grading is standardised using an image reference atlas. The following definitions are used:

Synovitis: “an area in the synovial compartment that shows above normal post-gadolinium enhancement of a thickness greater than the width of normal synovium.” Graded 0-3 based on proportion (in thirds) of enhancement of synovial tissue: 0 – normal, 1 – mild, 2 – moderate, 3 – severe

Bone erosions: “a sharply margined bone lesion, with correct juxta-articular localisation and typical signal characteristics, which is visible in two planes with a cortical break seen in at least one plane.” Graded 0-10 based on proportion (in centiles) of eroded tissue compared to total bone volume: 0 – normal, 1 – 1-10%, 2 - 11-20%, 3 – 21-30% etc

Bone marrow oedema: “a lesion within the trabecular bone, with ill defined margins and signal characteristics consistent with increased water content.” Graded 0-3 based on proportion of bone displaying oedema: 0 – no oedema, 1 – 1-33% oedematous, 2 – 34-66% oedematous, 3 – 67-100% oedematous.

For each region examined, total scores for each component can be calculated by summing the individual scores for each individual joint (Table 5). Changes in total scores, and for individual joints, between baseline and follow-up scans can be used to describe the rate of improvement / progression in each component at an overall, or individual joint level.

Validation studies with multiple independent readers have demonstrated that the RAMRIS system has high intra- and inter-reader agreement with high intra-class correlation coefficients for each of the individual components (median ICC: synovitis 0.69-0.90, bone erosion 0.73-0.91, bone marrow oedema 0.79-0.98) (250). The RAMRIS system has shown high sensitivity to change with

low smallest detectable differences (median SDD: synovitis 1.89, bone erosion 5.53, bone marrow oedema 3.18) and low minimal detectable changes (median MDC: synovitis 19.8, bone erosion 3.69, bone marrow oedema 7.07) demonstrated for each component (307). Furthermore, the RAMRIS system appears to identify a higher number of patients with erosive progression than standardised plain radiography scoring systems (307).

	<b>Wrist</b>	<b>2<sup>nd</sup> – 5<sup>th</sup> MCPj</b>	<b>Total (Combined)</b>
<b>Synovitis</b>	0-9	0 - 12	0 – 23
<b>Bone Erosion</b>	0 – 150	0 – 80	0 – 230
<b>Bone Marrow Oedema</b>	0 - 45	0 - 24	0 – 69

**Table 6:** Potential range of scores for RAMRIS components for wrist and 2<sup>nd</sup>-5<sup>th</sup> MCPj separately and combined

**Plain X-ray Hands and Feet:** The majority of RA interventional clinical trials have traditionally reported changes in plain x-ray findings as the primary radiological outcome. Acquiring plain radiography images is generally cheap, fast and facilities are usually available in most clinical settings. However, in early inflammatory arthritis there is an increasing recognition that plain radiography alone is a relatively insensitive method of detecting structural damage and progression since the frequency of plain radiographic erosions is relatively low (308). Hence, to maximise sensitivity MRI has been chosen as the primary radiological outcome measure; though plain radiographs of hands and feet will still be collected to allow comparison with previously published clinical trials.

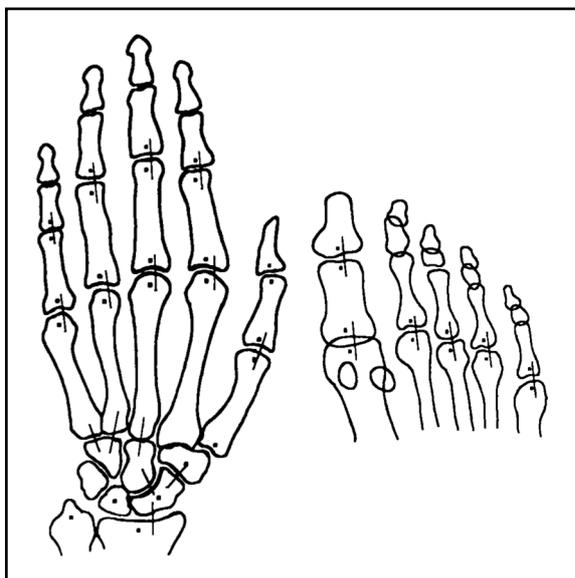
*The Sharp / Van der Heijde Score* (309): A recent survey of RA clinical trials identified the Sharp / Van Der Heijde Score as the most commonly reported plain radiographic outcome (310).

The presence and extent of erosions is assessed in the following joints: Hands – 1<sup>st</sup>-5<sup>th</sup> MCPj, 2<sup>nd</sup>-5<sup>th</sup> PIPj, IPj of thumb, 1<sup>st</sup> proximal metacarpal, distal radius and ulna, scaphoid, lunate, trapezium and trapezoid bones; Feet – 1<sup>st</sup>-5<sup>th</sup> MTPj and IPj of first toe. Erosions in the hands are graded 0-5: 0 – no erosion, 1 – single discrete interruption of cortical surface, 2-4 – erosive change involving between 2-4 quadrants of the joint, 5 – confluent erosions involving full surface of the joint. In the feet, each side of the joint is graded independently 0-5 and then both scores are summed to provide a total score out of 10 for each joint.

The presence of joint space narrowing is assessed in the following joints: Hands 1<sup>st</sup>-5<sup>th</sup> MCPj, 2<sup>nd</sup>-5<sup>th</sup> PiPj, 3<sup>rd</sup>-5<sup>th</sup> carpometacarpal (CMC), radiocarpal joint, scaphoid-lunate and lunate-capitate joints; Feet – 1<sup>st</sup>-5<sup>th</sup> MTPj and IPj of first toe. Joint space narrowing for hands and feet are graded 0-4: 0 – normal, 1 – focal or doubtful, 2 – involving ≤ 50% of joint surface, 3 – involving > 50% of joint surface or subluxation evident, 4 – bony ankylosis evident.

Grades for erosions and joint space narrowing can be summed for all joints giving a maximum erosion score of 160 in the hands, 120 in the feet and 280 altogether; a maximum joint space

narrowing score of 120 in hands, 48 in feet and 168 altogether. The maximum overall Sharp / Van der Heijde Score is 448. The Sharp / Van der Heijde system has demonstrated good metric properties with high scores for inter-rater reliability (ICC 0.80-0.96) and intra-rater reliability (ICC 0.94-0.99) for all components at single time points and in relation to detecting change (311,312). The smallest detectable difference between two radiographs is 7 (311).



**Figure 10:** Illustration of joints of hands and feet scored for erosions (•) and joint space narrowing (I) by the modified (Van der Heijde) Sharp Scoring System [Illustration from Van der Heijde et al. Arthritis Rheum 1992. 35. P26-34]

### ***Principles for Scoring of Radiological Analysis***

The following, over-arching principles will be applied to the grading of all plain radiography and MRI outcomes to ensure that the process is reliable and free from external bias.

1. Baseline and follow-up images at each site will be acquired using the same imaging equipment
2. Digital versions of all images will be reviewed and graded using the same image analysis platform ([www.osirix-viewer.com](http://www.osirix-viewer.com))
3. Each image will be reviewed independently by two musculoskeletal radiologists who are blinded to the participant's identity and treatment group.
4. For each image, the mean of the two independent gradings for each component will be used in the statistical analysis. However, if there is significant disagreement between the independent gradings for a particular image, the grading radiologists will be asked to discuss their assessments until a consensus is reached
5. Images will be reviewed in chronological order to increase the sensitivity to detect change over time (313)

#### **2.5.4 Data Storage and Management**

To facilitate the later statistical analysis, all data generated by this research will be stored within a secure, password-protected, online eCRF hosted and administered by the Robertson Centre for Biostatistics, University of Glasgow. Access to the eCRF will be restricted to those individuals responsible for assessing and recording each of the different data types. Each individual will only be able to enter data into the fields relating to their contribution and will not have sight of any

other outcome data. Furthermore, to maintain the objectivity and blinding of the outcome assessors, each patient's unique data entry pages will not display their randomisation group. Data queries will be administered by the Robertson Centre for Biostatistics whereby the relevant dataset will be downloaded from the 'live' eCRF since, to maintain data integrity, it will not be possible to analyse the data stored within the eCRF directly.

## **2.6 Statistical Analysis**

Due to the frequency and breadth of clinical and radiological data being collected a very large range of within group and between-group comparisons were possible. Comparisons were made to determine whether there was a significant difference in a single variable between both groups at a single time point or whether there had been a significant within group change over time of a particular variable. Prior to each analysis the distribution of the relevant dataset was determined. Given the group sizes it was presumed that most datasets would not be normally distributed and therefore, unless otherwise stated, non-parametric analysis techniques were used. Data for each comparison were drawn from the eCRF administered by the Robertson Centre for Biostatistics. All analyses were performed using either SPSSv17 or Graphpad Prism v6 ([www.graphpad.com](http://www.graphpad.com)). Unless otherwise stated,  $p < 0.05$  was considered the threshold for statistical significance.

### **2.6.1 Description of Baseline Characteristics**

In order to ensure that the control and intervention group could be directly compared, and that they were representative of a 'typical' RA population, it was important to determine whether there were any significant imbalances in baseline demographic or RA-related characteristics. Table 7 illustrates how each of the baseline variables were categorised, summarised and compared between the control and intervention groups. For the purposes of this illustration it is presumed that data will not be normally distributed and therefore, unless otherwise stated, data were analysed using non-parametric methods.

	<b>Variable</b>	<b>Categorisation</b>	<b>Summary Statistic (range)</b>	<b>Between Group Comparison Technique</b>
<b>Demographic Data</b>	<b>Sex</b>	Dichotomous	Frequency / Percentage	Chi-squared
	<b>Age</b>	Continuous	Median (IQR*)	Mann Whitney U
	<b>Weight / Height (BMI)</b>	Continuous	Median (IQR)	Mann Whitney U
	<b>Smoking Status</b> (current / ex / never)	Nominal	Frequency / Percentage	Chi-squared
<b>Disease Characteristics</b>	<b>Symptom Duration</b>	Continuous	Median (IQR)	Mann Whitney U
	<b>RhF Status</b> (positive or negative)	Dichotomous	Frequency / Percentage	Chi-squared
	<b>RhF Titre</b>	Continuous	Median (IQR)	Mann Whitney U
	<b>Anti-CCP Status</b> (positive or negative)	Dichotomous	Frequency / Percentage	Chi-squared
	<b>Anti-CCP Titre</b>	Continuous	Median (IQR)	Mann Whitney U
	<b>Baseline Erosions</b> (present / absent)	Dichotomous	Frequency / Percentage	Chi-squared
<b>Disease Activity / Impact</b>	<b>DAS44</b> (including its constituent parts)	Continuous	Median (IQR)	Mann Whitney U
	<b>HAQ</b>	Continuous	Median (IQR)	2 sample t test
	<b>CRP / ESR</b>	Continuous	Median (IQR)	Mann Whitney U
	<b>Pain 10cm VAS</b>	Continuous	Median (IQR)	Mann Whitney U

**Table 7:** Baseline characteristics and proposed statistical analysis methods (\*IQR – interquartile range)

### **2.6.2 Assessment of Impact of MSUS Upon DMARD Escalation Decision Making**

Since a central theme of this research is related to how MSUS findings could potentially influence DMARD escalation decisions it was important to describe how often there was agreement and disagreement between clinical and MSUS disease activity assessment. In this way, it was possible to determine how often MSUS findings would influence DMARD treatment decisions and also the potential impact of regularly incorporating MSUS into a rheumatologist's clinical work load. If there was a discrepancy in the assessment of disease activity, the MSUS findings took precedence over the DAS28 in determining whether or not DMARD escalation was indicated.

## ***Definitions of Agreement and Disagreement***

The following definitions of agreement and disagreement were used:

*Agreement:* DAS28 and MSUS both agree on the presence / absence of active inflammatory disease and lead to the same decision relating to DMARD therapy. This related to the following scenarios:

1. Presence of active disease requiring DMARD escalation: DAS28 >3.2 with minimal clinical synovitis (e.g 0-1 swollen joints) AND MSUS identifies grade 1 (or higher) PDUS signal in at least two joints
2. Absence of active disease requiring no change to DMARD therapy: DAS28 <3.2 AND MSUS identifies PDUS findings in either one or no joints

*Disagreement:* DAS28 and MSUS disagree on the presence / absence of active inflammatory disease and support opposing decisions relating to DMARD therapy. This related to the following scenarios:

1. Low disease activity but subclinical synovitis requiring DMARD escalation: DAS28 <3.2 BUT MSUS identifies grade 1 (or higher) PDUS signal in at least two joints
2. Moderate disease activity but absent synovitis on MSUS assessment: DAS28 >3.2 with minimal clinical synovitis (e.g 0-1 swollen joints) BUT MSUS identifies PDUS findings in either one or no joints

Since there was a three month gap between the decision to escalate DMARD therapy and the next opportunity to perform MSUS the amount of MSUS data available was limited. In order, to maximise the data available, all instances when there was paired sets of DAS28 and MSUS findings were pooled together. The following analyses were planned:

1. Description of percentage agreement and disagreement: using the definitions of agreement and disagreement described in the preceding section the frequency and percentage of agreement between DAS28 and MSUS was calculated. Since DAS28 and MSUS assess RA disease activity by quite different methods and return quite different outputs it was not possible to calculate a traditional statistical measure of agreement (e.g Kappa statistic). Thus, the simpler percentage agreement between the findings was chosen as the preferred descriptor
2. Frequency of joint involvement: the frequency of each joint displaying positive MSUS findings was calculated to determine the relative contribution of individual joint areas to the final disease activity assessment. Findings for corresponding right and left hand sided joints were pooled to calculate the frequency of positive findings in each area. Subsequently, whether one or both of a particular joint displayed positive findings was determined. In this way it was possible to identify joint areas that could be excluded from the proposed global MSUS assessment tool without negatively affecting the overall sensitivity

3. Joint scores and counts: at a single sitting gradings from each of the MSUS findings were combined in the following ways to provide an overall summary of the assessment:
  - i. Total MSUS Score – sum of all the gradings from each joint for a particular finding. Ranges: synovial hypertrophy 0-42, PDUS 0-42
  - ii. MSUS Joint Count – number of examined joints exhibiting a particular MSUS finding, regardless of the finding's grading. Ranges: synovial hypertrophy 0-14, PDUS 0-14. Patients exhibiting a PDUS index greater than or equal to 2 were eligible for DMARD escalation

For each finding, changes in mean MSUS score and indices over time were used to describe how each MSUS finding fluctuated over the follow-up period and responded to increasing DMARD intensity. Furthermore, longitudinal changes in MSUS findings were compared to corresponding changes in clinical disease activity assessment to determine how well the measured clinical changes reflected changes visible in underlying synovitis

### **2.6.3 Description of Treatment Intensity**

The specific details of each patient's DMARD regimen (e.g doses and constituent agents) and corticosteroid requirements was collected during each monthly consultation and recorded in the online eCRF

#### ***DMARD Therapy***

It was presumed that participants in the MSUS assessment group would receive more intensive DMARD therapy over the duration of the follow-up period; i.e more treatment escalation steps and a higher frequency, and earlier, use of combination DMARD and anti-TNF $\alpha$  blocking therapy. In order of ascending intensity, grading of treatment intensity was based upon the hierarchy suggested by the DMARD escalation protocol (section 2.3):

- |      |                            |  |
|------|----------------------------|--|
| i.   | DMARD Monotherapy          | Methotrexate, sulfasalazine or hydroxychloroquine  |
| iia. | Dual Combination Therapy   | Methotrexate + sulfasalazine; methotrexate + hydroxychloroquine; sulfasalazine + hydroxychloroquine      |
| iib. | Triple Combination Therapy | Methotrexate + sulfasalazine + hydroxychloroquine  |
| iii. | Subcutaneous methotrexate  | Subcutaneous methotrexate with one or both of sulfasalazine and/or hydroxychloroquine                    |
| iv.  | Biologic therapy           | Etanercept with some or all of methotrexate (oral or subcutaneous), sulfasalazine and hydroxychloroquine |

The proportion of patients in each assessment group within each of these treatment intensity groups was calculated for each 3 monthly time point (0, 3, 6, 9, 12, 15, 18 months). Furthermore, the mean dose of each individual DMARD being administered at each 3 monthly time point was calculated.

### ***Corticosteroid Treatment***

Since, corticosteroids have a rapid immunomodulatory action that can produce rapid, short-term fluctuations in measured disease activity it is feasible that unequal use of corticosteroids could account for some of the differences in outcome observed between each of the assessment groups. Whilst both assessment groups share the same indication for administering corticosteroids, it is feasible that an inadvertent treatment bias may be introduced which favours the MSUS assessment group. The intensity of corticosteroid treatment administered to each group was determined by:

- i. describing the mean number and cumulative dose of intra-articular and intra-muscular corticosteroids administered over the whole duration of the research,
- ii. describing the mean number and total dose of intra-articular and intra-muscular corticosteroids administered during each monthly review

Individual joints were injected with differing doses of triamcinolone acetonide depending upon their size; thus, for comparison purposes the mean cumulative dose of intra-articular and intra-muscular corticosteroid administered over the duration of the follow-up period will provide a more accurate representation of treatment intensity

#### **2.6.4 Description of Treatment Response**

A broad range of clinical and functional outcome measures were collected at 3 monthly intervals over the duration of the follow-up period and were in-line with the joint ACR-EULAR recommendations for reporting of clinical outcomes (314). This dense outcome data collection allowed each patient's, and their assessment groups, overall treatment response to be described by a variety of different measures. Numerical changes in composite disease activity measures - in particular the DAS44 – between two time-points allowed treatment responses to be described as either the absolute change in value or whether or not a predefined target (such as low disease activity or remission) was achieved. Furthermore, both the relative change from baseline of the DAS44 (e.g EULAR response criteria), and its mean over the follow-up period, were used as a measure of overall disease control to determine whether either group had been exposed to a significantly greater inflammatory burden. Broadly, changes in outcome measures between baseline and 3 months were presumed to represent changes in response to initial DMARD monotherapy; whereas, changes between baseline and 18 months were presumed to represent overall response to intensive DMARD therapy

#### ***DAS44***

The mean change in DAS44 from baseline was the primary clinical outcome measure. The numerical value of the DAS44 at each time point could also be manipulated by a variety of different additional methods to provide a measure of treatment response between different time-points and also over the duration of the follow-up period.

*Longitudinal change in DAS44:* It is commonly accepted that a fall in DAS44 of 1.2 (or greater) represents a significant, positive treatment response (297). The between group difference in the mean change in DAS44 between baseline and month 18 was compared to determine whether either group experienced a statistically higher magnitude of change in disease activity. The mean DAS44 for each group at each 3 month time point was compared using the 2 sample t-test test to determine if there were any statistically significant between group differences in disease activity at any time point

*Disease activity level thresholds:* at a given time point the absolute value of DAS44 allowed participant's disease activity to be categorised according to well established criteria (286,296):

Remission	DAS44 < 1.6
Low disease activity	DAS44 ≤ 2.4
Moderate disease activity	2.4 < DAS44 ≤ 3.7
High disease activity	DAS44 > 3.7

The proportion of each assessment group falling within each of the DAS44 categorisation groups at each time point was calculated and compared using a Chi-squared test

*EULAR response criteria:* between two time points, the absolute change in DAS44 and the final DAS44 values were used to determine the proportion of patients within each assessment group fulfilling the EULAR response criteria (297). EULAR responses were defined using the criteria described in Table 5 (reproduced below for ease)

	Improvement in DAS44		
DAS44 at endpoint	≤ 1.2	> 0.6 and ≤ 1.2	≤ 0.6
≤ 2.4	Good	Moderate	None
>2.4 and ≤ 3.7	Moderate	Moderate	None
>3.7	Moderate	None	None

**Table 5:** Derivation of EULAR response using DAS44 thresholds

The proportion of patients in each assessment group meeting each of the EULAR response criteria at each time point were compared using the Chi-squared test. There was particular focus on the EULAR responses between baseline and 3 months follow-up (response to initial DMARD monotherapy) and baseline and 18 months follow-up (overall response to intensive management).

*Cumulative inflammatory burden:* patient's response to DMARD therapy is neither uniform nor predictable and month-by-month fluctuations in DAS44 may not give a clear indication of a patient's overall exposure to active disease. For this study, each participant's mean DAS44 was presumed to provide a truer representation of their overall disease course and cumulative inflammatory disease burden. Therefore, each participant's mean-DAS44 period was calculated and then pooled within assessment groups. Median mean-DAS44 were compared between the assessment groups using the Mann-Whitney U test to determine whether or not there was a statistical difference in either group's overall exposure to active disease.

### ***Health Assessment Questionnaire***

The Likert responses to each of the HAQ questions was used to calculate a numerical HAQ score that represented a participants level of functional ability at each assessment time point (298,299). Between group comparisons of differences in functional ability were conducted by comparing each group's median HAQ using the Mann Whitney U test at each 3 month assessment time point

### ***Euro-QOL 5D-3L Questionnaire***

The 5 domains of the EQ-5D questionnaire rate a patient's ability to undertake specific descriptors of health status (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) using a 3 point Likert scale and a 20cm VAS (301). Whilst graded 1 to 3, the levels for each health status are cardinal variables with no numerical value. At each 3 monthly time point, each assessment group's health profile will be described by reporting the proportion of patients who fall within each level of the individual domains. Between assessment groups, the proportion of patient's within each level of individual domains will be compared using the Chi squared test and the median EQ-5D 20cm VAS will be compared using the Mann Whitney U test.

### ***Patient Global Health and Pain 10cm Visual Analogue Scales***

Ten centimetre VAS scales provide numerical representations of the patient's overall health perception and experience of pain in the preceding week. At each 3 monthly time point, median values for global health 10cm VAS and pain 10cm VAS were calculated and between group comparisons were performed using Mann Whitney U test

### ***Acute Phase Reactant Levels***

Values for CRP and ESR, corresponding to each of the 3 monthly assessment time points, were recorded from the laboratory result systems of NHS GGC. Median values of each reactant were calculated for each group at each of the 3 monthly assessment time points and between group comparisons at each time point were performed using the Mann Whitney U test. Mean and area under the curve values could also be used for each reactant to provide an additional measure of cumulative inflammatory burden.

### ***Composite Measure of Treatment Response***

The clinical and laboratory outcome data collected as part of this research comprised the ACR core set variables and therefore the degree of change in each core set variable over the duration of the follow-up period was used to determine the ACR response rate for each group (302,315). Each of the ACR response definitions (ACR20, ACR50, ACR70) require a minimum amount of improvement in both tender and swollen joint count and a minimum amount of improvement in 3 of the remaining 5 variables. The ACR core set variable comprise:

- |                              |                         |
|------------------------------|-------------------------|
| 1. Tender joint count        | Ritchie articular index |
| 2. Swollen joint count       | 44 swollen joint count  |
| 3. Patient pain assessment   | Pain 10cm VAS           |
| 4. Patient global assessment | Patient Global 10cm VAS |

5. Physician global assessment	Physician Global Likert Scale
6. Assessment of physical function	HAQ questionnaire
7. Acute phase reactant value	ESR

ACR response rates were calculated between baseline and 6 monthly time points using the following method

1. At least 20/50/70% improvement in tender joint count AND swollen joint count  
AND
2. At least 20/50/70% improvement in 3 out of the following 5 variables:
  - i. pain 10cm VAS
  - ii. patient global 10cm VAS
  - iii. physician global likert scale
  - iv. HAQ questionnaire
  - v. ESR

The proportion of patients within each assessment group meeting each of the ACR response definitions at each 6 monthly time point were compared using the Chi-squared test.

### **2.6.5 Description of Change in Radiological Outcomes**

Changes in radiological measures over the follow-up period provide an additional measure of treatment success. The group with the least effective assessment strategy will be expected to exhibit either a greater frequency and degree of erosive progression and a greater persistence of active synovitis. Changes in MRI appearances are considered the primary radiological outcome; however, changes in plain x-ray appearances will also be reported since these remain the most commonly reported radiological outcome in RA clinical trials. At the time of writing, the formal grading of the radiological outcomes had not been completed; therefore it will not be possible to present the radiological outcome results within this thesis

#### ***MRI Outcomes***

Baseline and 18 month MRI appearance of synovitis, erosions and bone marrow oedema will be graded using the previously described RAMRIS system (section 2.5.3). For each assessment group at each time point, median grades for each component will be reported for the wrist and MCPj individually and as a combined score. The mean change in the numerical value of the each component's grade between baseline and 18 months will represent the impact of the assessment strategy on disease progression and treatment response. The group with the most effective assessment and treatment strategy will be expected to demonstrate a lesser increase in the erosion score and a greater reduction in the synovitis and bone marrow oedema scores. Median scores at each time point, and their change over the follow-up, will be compared between the assessment groups using either the Mann Whitney U test or Student's t test as appropriate. Table 6 (reproduced below for ease) summarises the range of scores for each component by anatomical area

	<b>Wrist</b>	<b>2<sup>nd</sup> – 5<sup>th</sup> MCPj</b>	<b>Total (Combined)</b>
<b>Synovitis</b>	0-9	0 - 12	0 – 23
<b>Bone Erosion</b>	0 – 150	0 – 80	0 – 230
<b>Bone Marrow Oedema</b>	0 - 45	0 - 24	0 – 69

**Table 6 (reproduced):** Potential range of scores for RAMRIS components for wrist and 2<sup>nd</sup>-5<sup>th</sup> MCPj separately and combined

### ***Plain X-ray Outcomes***

Baseline and 18 month plain x-rays of the hands and feet will be graded using the Sharp / Van der Heijde Score (311). Mean changes between baseline and 18 months in the erosion, joint space narrowing and total scores will be used to represent the degree of radiographic progression evident in both groups. It is presumed that the group with the most effective assessment and treatment strategy will also exhibit a lesser increase in all components of the Sharp / Van der Heijde Scores. Median scores at each time point, and their mean change over the follow-up period, will be compared between assessment groups using either the Mann Whitney U test or Student's t test as appropriate. Table 8 summarises the potential range of scores for each anatomical region

	<b>Erosions</b>	<b>Joint Space Narrowing</b>	<b>Total</b>
<b>Hands</b>	0-160	0-120	0-280
<b>Feet</b>	0-120	0-48	0-168
<b>Total</b>	0-280	0-168	0-448

**Table 8:** Potential range of scores for plain xray erosions and joint space narrowing in the hands and feet when graded using the Sharp / Van der Heijde Score

To provide some indication of the burden of erosive disease at presentation the presence or absence of erosions on baseline hand and foot x-rays was recorded from the standard radiological reports issued by NHSGGC staff radiologists. These reports are provided by a wide number of radiologist with varying degrees of experience in reporting plain x-ray findings. Reports are descriptive and do not formally quantify the presence / absence of radiological features of RA. Consequently, the presence / absence of baseline x-ray erosions reported by this thesis is not standardised and subject to significant inter-reader variability. It is highly likely that the formal grading of plain x-ray images using the modified Sharp score will return findings that differ significantly from the values quoted by this thesis (especially the prevalence of erosive change at baseline)

### **2.6.6 Adverse Event Rates**

Even though the treatment protocol comprises several DMARDs that are commonly used in combination in the treatment of RA, it is possible that the aggressive escalation strategy favoured by the MSUS assessment group could lead to a higher incidence of adverse effects. In fact, it

might become evident that the frequency of adverse effects observed when patients with asymptomatic, subclinical synovitis receive increasingly aggressive DMARD combinations tips the risk:benefit ratio in favour of not escalating DMARD therapy. Equally, the prevention of DMARD escalation by MSUS findings excluding on going synovitis in patients with elevated disease activity scores might also divert some patients away from the risk of increased risk of adverse effects associated with combination therapy. Until the frequency of adverse effects has been determined it will not be possible to comment on the potential safety implications of using MSUS to guide DMARD escalation. Throughout the duration of the study, the incidence, duration and nature of adverse events was carefully recorded using standardised preforms provided by the Robertson Centre for Biostatistics, University of Glasgow. The definition of an adverse event is based upon the standardised definitions published by the Medicines and Healthcare products Regulatory Agency (MHRA):

*Adverse Event:* any untoward medical occurrence in a subject to whom a medicinal product has been administered, including occurrences which are not necessarily caused by or related to that product

*Adverse Drug Reaction:* any untoward and unintended response in a subject to an investigational medicinal product which is related to any dose administered to that subject

*Unexpected Adverse Reaction:* an adverse reaction the nature and severity of which is not consistent with the information about the medicinal product in question set out:

- i. in the case of a product with a marketing authorisation, in the summary of product characteristics for that product
- ii. in the case of any other investigational medicinal product, in the investigator's brochure relating to the trial in question

*Serious Adverse Event / Unexpected Serious Adverse Reaction:* any adverse event, adverse reaction or unexpected adverse reaction, respectively, that:

- i. results in death
- ii. is life threatening
- iii. requires hospitalisation or prolongation of existing hospitalisation
- iv. results in persistent or significant disability or incapacity
- v. consists of a congenital anomaly or birth defect

Important medical events that may not be immediately life-threatening or result in death or hospitalisation but may jeopardise the patient or may require intervention to prevent one of the other outcomes listed in the definition above are also considered to be serious

At the time of writing the adverse event data were not available for analysis. In due course, the total frequency of adverse events occurring within each assessment group will be reported for the

duration of the assessment period. Adverse events will be categorized according to severity (based on the preceding definitions) and also nature.

## 2.7 Biomarker Analysis

A nested biomarker study was conducted in parallel to the clinical study. All participants recruited to the clinical study donated additional blood samples at set time points throughout the follow-up period. The ultimate aim was that these samples would undergo analysis using a variety of different molecular platforms so that changes in biomarker signature could be compared to corresponding changes in clinical measures of disease activity and treatment response. This clinical data was already being collected as part of the clinical study. Any samples that were not immediately analysed were stored as a research tissue bank and made available for future exploratory analyses in relation to the research cohort.

### 2.7.1 Principles of Biomarker Analysis

#### *Collection and Storage of Samples*

Table 9 summarises which additional blood samples were collected and their intended use:

Draw Order	Vacutainer	Fraction	Quantity	Proposed Analysis
1	SST - Serum Separation Tube	Serum	1	Immunoassay
2	EDTA	Plasma	2	Proteomics
3	Lithium Heparin	Plasma	2	Immunoassay
4	BD P100	Plasma	1	Proteomics
5	PAXgene RNA	Whole blood	1	RNA
6	PAXgene DNA	Whole blood	1	DNA

**Table 9:** Vacutainer set and draw order

PAXgene RNA and DNA samples were stored in the original vacutainers whereas all other aliquots were stored in polypropylene tubes. All samples were labelled using each participant's unique study identifier and a code relating to the sample type and sampling time point. Table 10 summarises the sample labelling system for (T00X –  $\alpha$  –  $\beta$ ).

T00X	- $\alpha$	- $\beta$
Patient's unique study identifier Chronological 0-111	Sample time point	Sample type
	A – baseline	S – SST
	B – 3 months	E – EDTA
	C – prior to etanercept	L – Lithium Heparin
	D – 3 months of etanercept	P – BD P100
	E – 6 months of etanercept	RNA – PAXgene RNA
	F – 18 months (study completion)	DNA – PAXgene DNA

**Table 10:** Illustration of biomarker sample labelling system for sample (T00X –  $\alpha$  –  $\beta$ )

In order to reduce any errors that could have been introduced by variations in sampling handling procedures all initial handling and aliquoting was performed by Dr James Dale following a standardised procedure based on each vacutainer's manufacturer's instructions and the advice of laboratory colleagues based at the Translational Medicine Research Collaboration, Dundee. Even

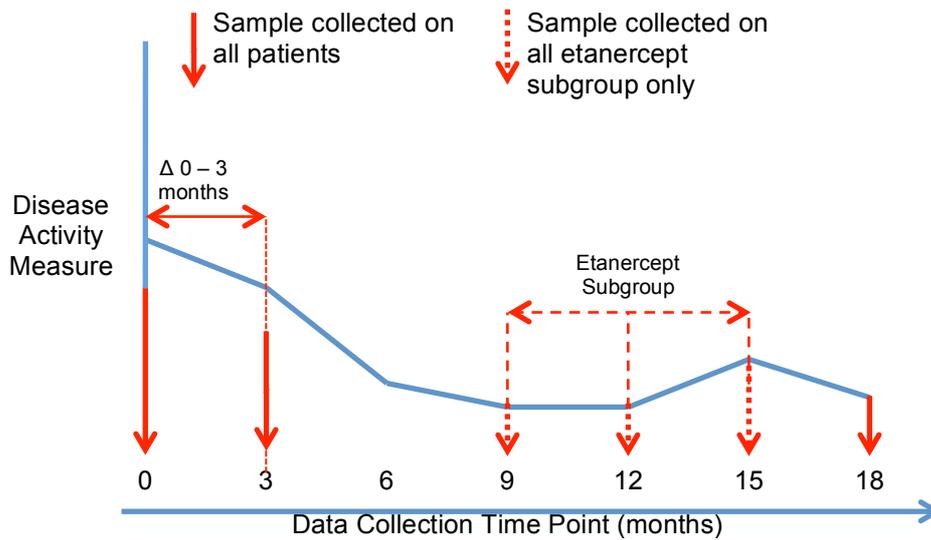
though the eventual biomarker analysis did not happen in collaboration with TMRC their recommendations for sample handling were continued for the duration of the study. Table 11 summarises the initial sample handling procedures and storage arrangements for the different sample types from the point of collection to first storage. The target was that all samples should be placed within the storage freezers within 4 hours of initial collection. Samples were initially stored together in the freezers of the Glasgow Biomedical Research Centre, University of Glasgow though the collection has recently been transferred to the freezers of NHSGCC Biorepository and catalogued using the Laboratory Information and Management System (LIMS).

Labelling Code	Vacutainer	Number of Inversions	Transfer Temperature	Centrifugation (1100g – 12 minutes)	Aliquots	Final Storage Temperature
<b>S</b>	SST	5	Ice	Yes	3-4 x 500µl Serum	-80°C
<b>E</b>	EDTA	8-10	Ice	Yes	5-7 x 500µl Plasma	-80°C
<b>Li</b>	Lithium Heparin	8-10	Ice	Yes	5-7 x 500µl Plasma	-80°C
<b>P</b>	BD P100	8-10	Ice	Yes	5-7 x 500µl Plasma	-80°C
<b>RNA</b>	PAXgene RNA	8-10	Room	No	1 x 2.5ml Whole blood	-80°C
<b>DNA</b>	PAXgene DNA	8-10	Room	No	1 x 8.5ml Whole blood	-80°C

**Table 11:** Initial blood sample handling procedures

### ***Sampling Time Points***

Figure 11 illustrates at which points during the follow-up period participants were asked to donate additional blood for biomarker analysis. All patients who completed the full follow-up period had donated sample sets at baseline, after 3 and then 18 months of participation. The small subset of patients who qualified for etanercept therapy were also asked to donate additional sample sets at the point they commenced etanercept, after 3 months of etanercept and after 6 months (completion of etanercept).



**Figure 11:** Illustration of relationship between sample set collection time points and changes in disease activity

The longitudinal clinical outcome data collection and the various sampling time points were chosen to explore how dynamic changes in clinical disease activity were reflected by changes in corresponding biomarker expression profiles. These timings allowed the following relationships to be explored:

1. Clinical disease activity measure at a single time point VS corresponding biomarker expression signature: to determine degree of correlation between clinical and biomarker measures of disease activity, and thereby to determine whether particular biomarkers might act as additional measures of global disease activity
2. Presenting phenotypic profile VS baseline biomarker signatures: to determine whether specific phenotypic groupings (e.g. rheumatoid factor status, anti-CCP antibody status) were associated with specific patterns of biomarker expression
3. Baseline biomarker expression signature VS longitudinal disease activity outcome data: to determine whether baseline biomarker signature associate with, and were therefore predictive of, specific patterns of either disease activity, treatment response (e.g. persistently active or inadequate treatment responders) and/or adverse events
4. Three month clinical disease activity measure VS three month biomarker expression profile. This is a strategically crucial point in a patient's treatment course since it marks the first time that the response to initial DMARD monotherapy, and therefore the need to possibly escalate DMARD therapy, is considered. At this point biomarker signatures could theoretically serve several overlapping, prognostic purposes: 1. identification of those who will achieve and sustain a good response to DMARD monotherapy (e.g. good prognosis); 2. identification of those who exhibit active disease and exceed DMARD escalation thresholds; 3. differentiation between patients with different classifications of disease activity (particularly the separation of patients with true remission from those with subclinical synovitis in the absence of clinically evident synovitis)

5. Change between two points of clinical disease activity measure VS corresponding change in biomarker expression signatures. In this way it might be possible to identify specific changes in biomarker signatures that are related to specific clinical responses following DMARD changes. Changes between baseline and 3 months represent response to initial DMARD monotherapy whereas changes between baseline and 18 months represent overall response to intensive step-up therapy

## 2.7.2 Transcriptomic Analysis

### Grateful Acknowledgement

The preparation and processing of the PAXgene RNA samples detailed in the following sections was performed with the kind assistance of Dr Martin McBride (lecturer), Dr Wai Kwong Lee (research manager), Mrs Iona Donnelly (research associate) and Miss Wendy Crawford (technician) of the Institute of Cardiovascular and Medical Sciences and Mrs Lynn Crawford (senior technician) of the Institute of Infection, Immunity and Inflammation.

The following description of sample handling, purification and analysis is based upon the standard techniques currently used by the Systems Biology Group of the Institute of Cardiovascular and Medical Sciences at the University of Glasgow. In turn, these procedures are largely based upon the recommendations of the manufacturers (Preactix, Qiagen Group, Becton, Dickinson and Company, Franklin Lakes, New Jersey, USA) of the PAXgene RNA collection and purification equipment. The PAXgene RNA system allows reliable storage and transportation of human whole blood and integrates with an efficient method of intracellular RNA purification. Each PAXgene RNA tube contains a proprietary mixture of RNA stabilization compounds that minimise RNA molecule degradation by RNases and ex vivo shifts in gene expression. For this research, all RNA purification steps were performed manually. The PAXgene RNA manufacturer's literature suggests that, on average, this technique yields at least 3µg of RNA from 2.5ml of whole blood on at least 95% of all samples. The necessary materials and equipment for conducting the analysis are listed in Appendix F

### 2.7.2.1 RNA Concentration and Purification

1. Baseline and 3 month PAXgene RNA samples identified, removed from -80 degree freezer storage and brought slowly back to ambient temperature. PAXgene RNA tubes were kept at ambient temperature for at least 2 hours prior to processing
2. Centrifugation of PAXgene RNA tube (10 minutes; 3000-5000g using swing-out rotor); Supernatant was decanted by pipette and nucleic acids contained within pellet were washed and re-suspended using 4ml RNase-free water
3. Protein digestion was triggered by incubating the re-suspended pellet with 40µl proteinase K, 350µl re-suspension and 300µl binding buffers. Incubated at 55°C for 10 minutes
4. Lysate was pipetted into PAXgene Shredder spin column and centrifuged for 3 minutes at maximum speed (up to 20,000g) to homogenise the cell lysate and remove residual cell debris.
5. 350µl ethanol (96-100% purity) was mixed with the supernatant.

6. 700µl of mixture was pipetted into the PAXgene RNA spin column and centrifuged at 8000-20,000g for 1 minute. RNA binds to the PAXgene silica membrane and contaminants were extracted by flow through
7. Repeated washes with wash buffer to remove any remaining contaminants. Between the first and second wash steps any residual bound DNA was removed by treating the silica membrane with DNase incubation mixture (10µl DNase I added to 70µl DNA digestion buffer). During each washing step the PAXgene RNA spin column was centrifuged (1 minute; 8000-20,000g) with 350-500µl wash buffer; contaminants were contained within the flow through liquid which was discarded
8. RNA was collected through elution by placing 40µl of elution buffer directly onto the PAXgene RNA spin column and centrifuging for 1 minute at 8000-20,000g. Elution solution was incubated at 65°C for 5 minutes to denature RNA

### **2.7.2.2 cDNA Hybridisation and RNA Amplification**

RNA purified from whole blood must be amplified, via a cDNA hybridisation step, to generate sufficient nucleic acid material to allow application to the microarray chip. An Illumina Beadchip microarray was used for this research; therefore, the standardised Illumina TotalPrep RNA Amplification procedure (316) was followed. The broad steps are:

#### *Reverse Transcription to Synthesise First Strand cDNA*

1. Nuclease-free water was added to 500 ng of total RNA to make a volume of 11µl
2. RNA solution was mixed with 9µl of Reverse Transcription Master Mix. Solution was centrifuged briefly to collect reaction at the bottom of the tube.
  - Reverse Transcription Master Mix comprises: 1µl T7 Oligo(dT) Primer, 2µl 10X First Strand Buffer, 4µl dNTP Mix, 1µl RNase Inhibitor, 1µl ArrayScript
3. Solution was incubated at 42°C for 2 hours in a thermal cycler and then briefly centrifuged

#### *Synthesis of Double-Stranded DNA from Single-Stranded cDNA by DNA polymerase*

4. 80µl of Second Strand Master Mix was added to RNA / Reverse Transcription Master Mix solution and mixed thoroughly
  - Second Strand Master Mix comprises: 63µl Nuclease-free Water, 10µl 10X Second Strand Buffer, 4µl dNTP Mix, 2µl DNA Polymerase, 1µ RNase HI
5. Solution was incubated at 16°C for 2 hours in a thermal cycler.

#### *Purification of cDNA*

6. 250µl of cDNA Binding Buffer was added to each double-stranded DNA sample and mixed thoroughly. Mixture was briefly centrifuged (10,000g) to collect reaction at the bottom of the tube
7. Double-stranded DNA mixture was pipetted onto the centre of the cDNA Filter Cartridge which was then centrifuged for 1 minute at 10,000g until the mixture has passed through the filter. All flow-through was discarded

8. 500 µl of wash buffer was added to cDNA Filter Cartridge and the mixture was centrifuged again at 10,000g for 1 minute until all the wash buffer had passed through the filter. All flow-through was discarded
9. 20 µl of preheated (50-55°C) Nuclease-free water was applied to the centre of the cDNA Filter Cartridge.
10. The cartridge was allowed to stand, at room temperature, for 2 minutes and was then centrifuged at 10,000g for 1 minute until all the Nuclease-free water had passed through the filter. The double-stranded DNA was present in the eluate (approximately 17.5µl)

#### *In Vitro Synthesis of cRNA from cDNA Templates*

11. 7.5 µl of IVT Master Mix was added to each double-stranded DNA sample, mixed and then incubated in a thermal cycler at 37 °C for 4-14 hours
  - IVT Master Mix comprises: 2.5 µl T7 10X reaction buffer, 2.5µl T7 enzyme mix, 2.5µl Biotin-NTP mix
12. 75µl of Nuclease-free water was added to the cRNA mixture to halt the reaction. The total volume was 100µl

#### *Purification of cRNA*

13. 350µl of cRNA Binding Buffer and 250µl of ACS reagent grade 100% ethanol was added to each RNA sample
14. cRNA – ethanol mixture was pipetted onto the centre of the cRNA Filter Cartridge which was centrifuged for 1 minute at 10,000g until all the mixture had passed through the filter. All the flow-through liquid was discarded
15. 650µl wash buffer was applied to each cRNA Filter Cartridge, centrifuged for 1 minute at 10,000g until all the wash buffer had passed through the filter. All the flow-through liquid was discarded
16. 200µl of preheated (50-55°C) Nuclease-free water was added to the cRNA solution and the whole mixture was incubated at 55°C for 10 minutes
17. The cRNA-water solution and cRNA Filter Cartridge were centrifuged at 10,000g for 1.5 minutes until all the water had passed through the filter. The cRNA had eluted into the Nuclease free water

#### *Quality Control Analysis of cRNA*

RNA quantity was measured using Nanodrop 1000 spectrophotometry which determined that the mean concentration of cRNA synthesised per sample was 357.8ng/ul (SD ±100.6). Each sample was measured in either duplicate or triplicate (depending on the level of agreement between the first two samples) and then averaged prior to dilution to 45.45ng/µl.

RNA quality was measured by passing all samples through an Agilent Technologies Bioanalyser. All samples achieved an RNA Integrity Number (RIN) of 6 or greater (range 6-9.5). Indeed, only 1 sample returned a RIN value less than the accepted threshold of seven and 6 samples returned RIN values between 7.0 and 7.5. Hence, due to this extremely small number, and that there was only one opportunity to extract RNA, all samples were allowed to continue in the study

### **2.7.2.3 *Illumina Beadchip Microarray Analysis***

Purified cRNA was hybridised onto commercially available Illumina HumanHT-12v4.0 Beadchip microarray chips (Illumina Inc, San Diego, California, USA). Each microarray chip allows simultaneous analysis of 12 separate samples. Broadly, each microarray well comprises microscopic polystyrene beads with each bead being associated with a specific genetic probe. Approximately, 25,400 genes are distributed over 47,321 probe positions, providing whole genome coverage. When purified cRNA is exposed to the microarray chip surface, individual cRNA fragments bind covalently to the complementary bead-bound genetic probe. Fluorescently labelled target probes are then added to determine the relative strength and frequency of binding between probe and cRNA fragment. The strength of the fluorescent signal emanating from each probe is dictated by the number of cRNA fragments binding to the probe and thus is proportional to the quantity of RNA present within the sample under examination.

Standardised sample handling and processing procedures were followed throughout. Individual samples were allocated to specific chips and wells using the following blocking hierarchy: randomisation group – sex – anti-CCP status – RhF status – smoking status – baseline xray erosions – 3 month DAS28 – 12 month DAS28 – first DMARD. This blocking technique attempted to minimise any additional variations in the observed expression profile by ensuring that other common demographic and disease-related factors that could also feasibly influence expression profile were evenly distributed throughout the microarray chips. All microarray chips were analysed using a Beadarray reader

#### *Hybridisation of RNA*

18. cRNA samples were heated to 65°C for 5 minutes, pulse centrifuged (250g) and allowed to cool to room temperature. Hybridisation buffer and humidity control buffers were heated to 58°C for 10 minutes and then also allowed to cool to room temperature. Beadchip microarrays were allowed to equilibrate with room temperature
19. 750ng of cRNA was pipetted into each hybridisation tube. 5µl RNase-free water was pipetted into cRNA sample tube with 10µl of hybridisation control buffer
20. The hybridisation chambers were assembled by fitting the Beadchip hybridisation chamber, chamber gasket and chamber insert together. 200µl of humidity control buffer was pipetted into each of the hybridisation chambers
21. Each Beadchip microarray was fitted into its hybridisation chamber insert. 15µl of cRNA sample was pipetted onto the centre of each inlet port to ensure that all sections of the stripe were covered
22. Sample laden Beadchip microarrays were loaded into each hybridisation chamber which was then sealed and incubated at 58°C for at least 14 hours (i.e overnight)

#### *Washing of Beadchip Micorarrays*

23. Hybridisation chambers were removed from the incubator and individual Beadchip micorarrays were submerged face-up in a beaker containing 3ml E1BC buffer diluted with 1L RNase-free water. The Beadchip microarray's coverseal was removed whilst still

- submerged in the E1BC buffer solution and Beadchip microarrays were transferred to a staining dish containing 250ml of Wash E1BC solution
24. High temperature wash: Beadchip microarrays were incubated with High-Temperature Wash Buffer in a Hybex waterbath for 10 minutes
  25. First room temperature wash: using a slide rack handle, Beadchip microarrays were plunged in-and-out of a staining dish containing 250ml Wash E1BC solution 10 times. The staining dish and Beadchip microarray were then shaken at a medium-low setting on an orbital shaker for 5 minutes
  26. Ethanol wash: Beadchip microarray was plunged in-and-out of a staining dish containing 250ml 100% ethanol 10 times. The staining dish and Beadchip microarray were then shaken at medium-low setting on an orbital shaker for 10 minutes
  27. Second room temperature wash: using a slide rack handle, Beadchip microarrays were plunged in-an-out of a staining dish containing 250ml Wash E1BC solution 10 times. The staining dish and Beadchip microarray were then shaken at a medium-low setting on an orbital shaker for 2 minutes
  28. Block: a Beadchip wash tray was placed on a rocker mixer and filled with 4ml Block E1 buffer. Beadchip microarrays were transferred into individual wash trays, coated with Block E1 buffer and rocked at medium speed for 10 minutes

#### *Detection of Signal*

29. 2ml of Block E1 buffer was mixed with 1:1,000 dilution of Cy3-Streptavidin and added to a Beadchip wash tray. The Beadchip was placed in the wash tray, covered and placed on a rocker-mixer, at medium setting, for 10 minutes
30. Third room temperature wash: using a slide rack handle, Beadchip microarrays were plunged in-an-out of a staining dish containing 250ml Wash E1BC solution 5 times. The staining dish and Beadchip microarray were then shaken at a medium-low setting on an orbital shaker for 5 minutes
31. After washing, racks of Beadchips were dried by immediately centrifuging them at 1400 rpm at room temperature for 4 minutes
32. BeadArray Reader Assessment: Beadchips were placed into the BeadArray Reader tray, registered using each Beadchip's unique barcode. Eight Beadchip's were imaged during each run of the BeadArray Reader. Beadchips were scanned using the BeadArray Reader's standard scanning protocol. During each run a lazer is shone across the Beadchips to excite the fluor of the hybridized single-stranded product attached to each bead. Light emissions from each fluor are recorded as high-resolution images which correspond to individual Beadchip sections

#### **2.7.2.4 Data Processing and Analysis**

##### **Grateful Acknowledgement**

The data handling, quality control analysis and analysis of gene expression levels was performed with kind advice and assistance from Dr John McClure, Lecturer at the Institute of Cardiovascular and Medical Sciences, University of Glasgow

All data handling and comparisons were conducted using R ([www.r-project.org](http://www.r-project.org)) a freely available, text based statistical analysis and computation environment. Certain steps within the analysis pipeline are performed using additional, open-source R add-on packages

#### *Pre-comparison Data Handling*

1. Quantile Normalisation: In order to fulfil the accepted presumption that most genes are neutrally expressed, with approximately similar distributions, and to eliminate systematic (i.e non biological and artefactual) variability, the raw BeadArray data underwent quantile normalisation using beadarray, a R add-on package ([www.bioconductor.org/package/2.11/bioc/html/beadarray](http://www.bioconductor.org/package/2.11/bioc/html/beadarray)) designed for the pre-processing and analysis of Illumina BeadArray raw data
2. ComBat Transformation: Principle component analyses demonstrated that the raw Illumina BeadArray data clustered into two distinct groups that related to the microarray chip batch. In order to minimise the potential influence of this evident batch effect on the variability of the observed results the data underwent ComBat transformation using the Surrogate Variable Analysis (sva) R add-on package (317). This method of reducing batch effect by using surrogate variables has been shown to reduce error rates and improve reproducibility in differential expression analysis experiments
3. Differential Gene Expression Analysis: Each of the following between group comparisons were conducted using the R add-on package Limma (Linear Models for Microarray Data – [www.bioconductor.org/packages/2.12/bioc/html/limma](http://www.bioconductor.org/packages/2.12/bioc/html/limma)) (318). Limma provides a suite of functions that are widely used to interrogate microarray data and assess differential gene expression between pre-determined groups. Its analysis methods are based along empirical Bayesian principles and correct for wide ranges in gene expression variability, even after normalisation steps have been performed. This approach limits the chance of generating false positive findings through multiple testing methods

#### *Identifying Differences of Gene Expression Between Different Phenotypic Groups*

Simply assuming that all instances of statistically significant differences in expression levels (i.e. the fold change) of particular genes are disease related may exclude biologically important genes and doesn't allow for artefactual outliers. Also, given the large number of genes on each microarray natural variability will make it likely that some genes will by chance exhibit statistically significant differences of expression between the comparator groups. Thus, between group comparisons performed by Limma consider both fold change and variability to generate a p value which estimates how likely the results could have occurred by chance alone. Further, calculating the false discovery rate corrects the p value in proportion to the number of multiple tests performed and thus limits the likelihood of identifying false positives (319). For this analysis, the adjusted p value significance threshold for each comparison was determined on an individual by-comparison basis. The target threshold was an adjusted p value less than 0.05; however, this was adjusted upwards for some comparisons depending upon their initial findings

### **2.7.2.5 Groupings for Comparison**

Multiple data and blood sample collection time points, the continuous nature of the DAS28 score - and the variety of different thresholds and formulae that can be used to interpret static and dynamic changes in the DAS28 value – could feasibly allow many different types of comparisons between the mRNA expression profile data and the available clinical outcomes. However, it is likely that a number of these comparisons would be either clinically irrelevant or pathogenetically tenuous. In order that the findings remain clinically relevant, expression profiles were subdivided into comparison groups that referred to commonly encountered phenotypes and clinical scenarios (i.e supervised grouping). For each comparison, participants were segregated into different groups using the demographic and/or clinical outcome data that was collected as part of their participation in the main clinical trial. Differences between each group's pooled gene expression profiles were then examined using the Limma package. When participants were segregated into two groups, direct comparisons of mRNA expression profiles were performed. Static comparisons between corresponding mRNA expression profiles and phenotypic groupings were used to determine whether clinically recognisable, phenotypic groupings were related to differences in gene expression. Dynamic comparisons that correlated gene expression profiles to a subsequent predefined clinical state (or change in state) were used to determine whether specific gene expression profile patterns might discriminate between – and therefore be predictive of – future clinical state or treatment response.

The DAS28 forms a continuous, numerical measure of disease activity. Whilst, there are established thresholds that differentiate between disease activity states, the threshold values do not relate to specific pathogenetic subsets of RA. Further, it is unlikely that there will be significant differences in gene expression evident in the profiles of phenotypically similar patients who fall either side of a DAS28 threshold. Therefore, comparator groups were created that exhibited large phenotypic differences since it was presumed that large phenotypic differences would most be most likely to be associated with evident differences in gene expression. In most cases, this method lead to comparator groups comprising the upper and lower quartiles of DAS28 (or its interpolations) at a particular time point. Prior to conducting gene expression analyses, the degree of difference in the phenotypic descriptor of each group (e.g median DAS28) was compared using non-parametric methods.

The following phenotypic grouping structures, and PAXgene RNA sampling time point comparator, were used to segregate mRNA expression profile data into distinct groups for comparison:

*Baseline phenotypic groupings:* groups defined using baseline demographic and RA-related phenotypic data. Baseline gene expression profiles served as the comparator set.

- Sex
  - 1. Female
  - 2. Male
  
- Smoking status
  - 1. Current smoker
  - 2. Former smoker

- 3. Never smoked
- Rheumatoid factor status
  1. Rheumatoid factor positive (titre >10iu/l)
  2. Rheumatoid factor negative (titre ≤10iu/l)
- Anti-CCP Antibody status
  1. Anti-CCP antibody positive (titre >10iu/l)
  2. Anti-CCP antibody negative (titre ≤10iu/l)
- Baseline plain X-ray status
  1. Erosive baseline x-rays
  2. Non-erosive baseline x-rays

*Disease activity at a single time point:* defined using existing DAS28 based definitions of disease activity. For each time point the corresponding gene expression profiles were used as the comparator set

- Baseline disease activity – quartiles based
  1. Highest baseline DAS28 quartile
  2. Lowest baseline DAS28 quartile (2<sup>nd</sup> and 3<sup>rd</sup> quartiles are excluded)
- 3 month disease activity – quartiles based
  1. Highest baseline DAS28 quartile
  2. Lowest baseline DAS28 quartile (2<sup>nd</sup> and 3<sup>rd</sup> quartiles are excluded)
- 3 month disease activity – threshold based
  1. High - DAS28 ≥ 5.1
  2. Moderate – 3.2 ≤ DAS28 < 5.1
  3. Low – 2.6 ≤ DAS28 < 3.2
  4. Remission – DAS28 < 2.6
- 3 month DAS28 < 3.2  
(Strategic threshold for DMARD escalation)
  1. 3 month DAS28 < 3.2
  2. 3 month DAS28 ≥ 3.2

*Change in disease activity over time:* defined using existing definitions of treatment response based upon the change in DAS28 between two time points. Both dynamic and predictive analyses were conducted and the gene expression profile comparator will be listed with the description of each comparison

Dynamic comparisons compare change in disease activity measure to corresponding change in gene expression profile

- Measured change in expression profile vs observed change in DAS28 – change in both variables between baseline and 3 months
  1. Lowest quartile change DAS28 0-3 months
  2. Highest quartile change DAS28 0-3 months (2<sup>nd</sup> and 3<sup>rd</sup> quartiles are excluded)

Predictive comparisons determine whether gene expression profiles at a single time point associate with particular patterns of subsequent clinical response

- Baseline profile vs 3 month DAS28  
(profile as predictor of need for DMARD escalation after 3 months monotherapy)
  1. 3 month DAS28 < 3.2
  2. 3 month DAS28 ≥ 3.2

- Baseline profile vs mean DAS28 0-18 months (profile as predictor of disease course and over all treatment response)
  1. Lowest quartile mean DAS28
  2. Highest quartile mean DAS28

*Overall treatment response:* the grouping definitions that underpin the analyses described in the preceding sections are based upon phenotypic variables from single time point and should lead to all participants being classified into one or other of the comparator groups. However, an individual patient's treatment response or disease character may not be evident from either a single, static measure of disease activity nor a measure of change between two relatively closely associated time points. In many cases, the character of a patient's RA will not become apparent until several months of therapy have elapsed by which time either the therapeutic window of opportunity may have passed and/or irreversible joint damage may have already occurred. It is presumed those patients who respond well to relatively light DMARD therapy may exhibit gene expression profiles that distinguish them from poorer prognosis patients who continue to display persistently active disease. Thus, the longitudinal clinical outcome data between baseline and 18 months will be used to categorise patients into groups based upon their overall treatment response. Those with extremes of response (e.g persistently active and persistent remission) will be compared to determine if there are any preceding differences in baseline gene expression profiles. The following grouping definitions will be used and all comparisons will be between the associated baseline mRNA expression profiles:

- Mean DAS28 0-18 months – quartiles based
  1. Worst treatment response  
Upper quartile mean DAS28 0-18
  2. Best treatment response  
Lower quartile mean DAS28 0-18
- Mean DAS28 0-18 months – criteria based (thresholds are arbitrary)
  1. Persistently active RA  
Mean DAS28 0-18 months > 4.2
  2. Persistent low disease activity  
Mean DAS28 0-18 months < 3.2

As previously described (Section 1.4.1), imaging studies consistently suggest that DAS28 remission (e.g. DAS28 < 2.6) don't necessarily equate with total absence of inflammatory disease activity. Rather, the level of active synovitis has simply fallen below a threshold that is no longer clinically detectable. This understanding that current remission definitions don't necessarily represent inactive disease has led to the development of much stricter (Boolean) remission criteria which require the virtual absence of all clinical measures of RA disease activity (e.g tender joint count  $\leq 1$ , swollen joint count  $\leq 1$ , CRP  $\leq 10\text{mg/l}$  and patient global 10 cm VAS  $\leq 1$ ) (186). Interestingly, and somewhat counter-intuitively, patients can return DAS28 scores upto 2.8 and yet still fulfil the Boolean definition of remission. Whether fulfilment of the Boolean remission criteria equates to total absence of inflammatory disease activity remains to be proven by further imaging studies; however it is still likely to be a truer approximation, given that patients can return DAS28 < 2.6 whilst still displaying 3-4 clinically swollen joints and/or a moderately elevated ESR/CRP. In reality, increasingly stricter definitions of remission may represent further steps along a continuum from clinically detectable active RA, to subclinical RA that can be

demonstrated by MSUS to absence of both clinical and imaging evidence of active synovitis. It is possible that the proposed Boolean definition of remission may still not be either sensitive nor specific enough to truly identify all patients with inactive RA and may yet overlap somewhat with existing definitions of remission. However, since it is the tightest clinical definition of RA remission currently available, it is the most efficient means of identifying those patients at the very lowest end of the disease activity continuum. Hence, it will be used as an alternative phenotypic definition in the following analyses to be conducted using the 3 month expression profile datasets:

- Differences between definitions of remission
  - 1. 3 month DAS28 < 2.6 but NOT Boolean remission
  - 2. Boolean remission at 3 months
  
- Extremes of disease activity at 3 months
  - 1. 3 month DAS28 > 4.6 and 2 or more swollen joints
  - 2. Boolean remission at 3 months

Grouping Variable	mRNA Expression Profile Timepoint	Groups
<b>Baseline Phenotypic Groupings</b>		
Sex	Baseline	1. Male 2. Female
Smoking Status	Baseline	1. Current smoker 2. Ex-smoker 3. Never smoked
Rheumatoid Factor Status	Baseline	1. RhF positive (titre >10iu/l) 2. RhF negative (titre ≤10iu/l)
CCP Antibody Status	Baseline	1. aCCP positive (titre >10iu/l) 2. aCCP negative (titre ≤10iu/l)
Baseline Erosions	Baseline	1. Erosions baseline x-rays 2. Non-erosive baseline x-rays
<b>Clinical Disease Activity – Single Time Point</b>		
Baseline Disease Activity - quartiles	Baseline	1. Lowest quartile baseline DAS28 2. Highest quartile baseline DAS28 (Middle 2 <sup>nd</sup> and 3 <sup>rd</sup> quartiles excluded)
3 month Disease Activity - quartiles	3 months	1. Lowest quartile 3 month DAS28 2. Highest quartile 3 month DAS28 (Middle 2 <sup>nd</sup> and 3 <sup>rd</sup> quartiles excluded)
3 month DAS28 Disease Activity - thresholds	3 months	1. High - DAS28 ≥ 5.1 2. Moderate – 3.2≤DAS28<5.1 3. Low – 2.6≤DAS28<3.2 4. Remission – DAS28<2.6
3 month DAS28 < 3.2	3 months	1. 3 month DAS28 < 3.2 2. 3 month DAS28 ≥ 3.2

**Table 12:** Summary of phenotypic grouping structures and corresponding mRNA expression profile sampling time points

Grouping Variable	mRNA Expression Profile Timepoint	Groups
<b>Change in Clinical Disease Activity</b>		
Change DAS28 0-3 months	$\Delta$ Expression profile 0-3 months	<ol style="list-style-type: none"> <li>1. Lowest quartile change DAS28 0-3 months</li> <li>2. Highest quartile change DAS28 0-3 months (2<sup>nd</sup> and 3<sup>rd</sup> quartiles are excluded)</li> </ol>
<b>Predictive properties of Baseline Expression Profile</b>		
DMARD escalation at 3 months	Baseline	<ol style="list-style-type: none"> <li>1. 3 month DAS28 &lt; 3.2</li> <li>2. 3 month DAS28 <math>\geq</math> 3.2</li> </ol>
Overall treatment response Mean DAS28 0-18 months - quartiles	Baseline	<ol style="list-style-type: none"> <li>1. Lowest quartile mean DAS28 0-18 months</li> <li>2. Highest quartile mean DAS28 0-18 months</li> </ol>
Overall treatment response Mean DAS28 0-18 months – criteria	Baseline	<ol style="list-style-type: none"> <li>1. Persistently active – mean DAS28 &gt; 4.2</li> <li>2. Persistent low disease activity – mean DAS28 &lt; 3.2</li> <li>3. Persistent clinical remission – DAS28 &lt; 2.6 on at least 6 occasions and meand DAS28 &lt; 4.2</li> </ol>
<b>Comparisons Using Boolean Definition of Clinical Remission</b>		
Alternative definitions of clinical remission	3 months	<ol style="list-style-type: none"> <li>1. 3 month DAS28 &lt; 2.6 but NOT Boolean remission</li> <li>2. Boolean remission at 3 months</li> </ol>
Extremes of disease activity	3 months	<ol style="list-style-type: none"> <li>1. 3 months DAS28 &gt; 4.6 and 2 or more swollen joints</li> <li>2. Boolean remission at 3 months</li> </ol>

**Table 12:** Summary of phenotypic grouping structures and corresponding mRNA expression profile sampling time points

### **3. Practical Considerations, Unexpected Developments and Research Design Amendments**

The methods proposed in preceding sections 2.1 to 2.7 describe the ideal means of addressing the clinical and transcriptomic research hypotheses. However, due to a number of unexpected logistical challenges it has not been possible to complete the full outcomes analysis and it will not be possible to present the complete results set. Rather, the results chapters of this thesis will describe as much of the analyses of the clinical, ultrasonographic and transcriptomic datasets as it was possible to present at this time. The contents of these analyses will be described in more detail at the start of each of the relevant chapters and the following sections will describe the sequences of events that led to this change in analysis plan

### **3.1 Dates**

The clinical and transcriptomic research projects described within this submission were first proposed in 2008 and the first protocol outline was completed in September 2008. However, by the time that the appropriate funding and approvals had been confirmed and the necessary logistical arrangements had been made (see below), it was not possible to recruit the first participant until September 2009. The final, 111<sup>th</sup>, participant was recruited in November 2011; therefore, the participant follow-up period ran until May 2013.

As the Chief Investigator, Dr Dale was the sole clinician responsible for the on-going clinical care of all the participating patients. In order to have sufficient time to screen, recruit and review all of the research participants, Dr Dale needed to 1. apply for Out-of-Programme Time for Research from NHS Education Scotland and the Joint Royal Colleges Post-Graduate Training Board's Rheumatology Speciality Advisory Committee and 2. secure fellowship funding for the duration of this time. Dr Dale was awarded a 3 year Chief Scientist's Office Clinical Academic Fellowship in September 2008 and took up this post (held at the University of Glasgow) on 1<sup>st</sup> April 2009 (the latest possible date). However, since recruitment couldn't commence until September 2009, and didn't complete until November 2011, the complete follow-up of all the study participants fell outwith the research time offered by the fellowship. Dr Dale completed the fellowship in April 2012, and returned to full clinical training, initially as an NHS-funded Specialist Registrar and then as a University of Glasgow funded Clinical Lecturer (post commenced June 2012). Thus, from April 2012 until the completion of the follow-up period, Dr Dale will have been continuing to review research participants, conduct analyses and draft this submission whilst also working in a full time clinical training post.

Unexpected delays in recruitment and the conduct of the transcriptomic analysis have also impacted upon the submission of this thesis. The original intention was that this thesis would detail the full sets of results relating to both the clinical and transcriptomic analyses. However, it became apparent that the follow-up period would extend beyond the submission deadline for the thesis and it did not appear possible to present the complete results set. Therefore, the interim analyses described in Chapter 7 were conducted. Furthermore, considerable delays were encountered finalising the specific arrangements for the transcriptomic analysis (described in section 3.8.2 below). The relevant samples did not undergo laboratory processing until June 2012

and the subsequent quality control and clinical-correlation analyses were conducted between July 2012 and December 2012.

## 3.2 Research Funding

Research funding was provided from a variety of different sources:

- Chief Scientist's Office Clinical Academic Fellowship (ref CAF / 08 / 03): total value £174,422. Awarded to Dr Dale to cover all salary, superannuation, degree registration, clinical trial regulation, training and some consumable costs incurred over the duration of Dr Dale's Out of Programme Time for Research
- Investigator Initiated Research Project Grant: awarded by Wyeth Pharmaceuticals to Drs Dale and Porter and Professor McInnes. Total value £197,200; to cover all additional NHS costs incurred through the conduct of this research; including, funding of research nurse's salary for the duration of the study, the use of NHS MRI and plain x-ray facilities and the use of NHS clinical spaces. Furthermore, Wyeth Pharmaceuticals undertook to provide a 6 month supply of etanercept to all study participants who qualified for biologic therapy as part of the treatment protocol
- Investigator Initiated Research Project Grant: awarded by Pfizer UK to Drs Dale and Porter and Professor McInnes. Total value £75,000; to cover the costs incurred by the proposed biomarker / transcriptomic analysis

## 3.3 NHS Research Ethics and Management Approvals

This study was conducted with the full approval of the West of Scotland Research Ethics Service and NHSGGC Research and Development department. However, whilst the applications for these approvals were commenced well before Dr Dale took up his research fellowship, delays in the granting of the final approvals led to significant delays in the commencement of recruitment

*West of Scotland Research Ethics Committee Approval (Ref: 09 / S0709 / 38):* The initial research ethics application was submitted on 1<sup>st</sup> April 2009 and an initial favourable opinion was granted on 16<sup>th</sup> June 2009. However, this opinion was based upon a randomisation process that minimised using anti-CCP antibody status as a discriminating variable. The research investigators were then informed that, due to funding issues, NHSGGC Immunology Laboratory were unable to guarantee that the anti-CCP assay would be available for the duration of the recruitment period. Thus, in order to avoid future randomisation irregularities it was decided that randomisation should be minimised based upon each participant's RhF status instead, since this test had been routinely available for many years. Whilst, switching to using RhF status to inform the minimisation process would have no impact upon what treatments a participant would receive

during the research it could lead to being allocated to a different assessment group and thereby, indirectly, a different treatment path. Hence, since this relatively minor change could potentially influence a patient's randomisation allocation the research ethics committee requested that a substantial amendment be submitted. A second favourable opinion was granted on 4<sup>th</sup> September 2009.

Following the initial favourable opinion, two additional substantial amendments to the original research protocols were submitted. The first (submitted January 2011) proposed bringing clinical and radiological outcome data collection in line with the newly formed Scottish Early Rheumatoid Arthritis Inception Cohort (see section 3.7 below) and that all tissue samples donated for research be stored within its tissue bank. The second (submitted August 2012) proposed that two 500µl of serum from each patient, at each sampling time point, be sent to Crescendo Biosciences (San Francisco, USA) to determine the degree of agreement between MSUS and a novel cytokine profile (Multibiomarker Based Disease Activity Test - MBDA) assessment of global disease activity.

*NHSGGC Research and Development Approval (Ref: GN09RH196):* The application for NHSGGC R+D approval was first submitted on 1<sup>st</sup> April 2009. The financial funding provided by Pfizer UK and the Chief Scientist's Office was paid directly to the University of Glasgow. However, since the participants for the research were identified and reviewed in NHSGGC out-patient clinics and day wards the NHSGGC R+D Office acted as the main research sponsor. Thus a three-way agreement was negotiated between NHSGGC R+D Office and the legal departments of the University of Glasgow and Pfizer UK. A separate Letter of Understanding was agreed between the NHSGGC R+D Office and the University of Glasgow that allowed the NHSGGC R+D Office to act as the main research sponsor whilst the University of Glasgow Finance Office administered the financial accounts (including payment to NHSGGC for the use of its facilities). The series of negotiations that were required to finalise the funding contracts created a significant temporal obstacle that prevented the commencement of participant screening and recruitment. Furthermore, Dr Dale and the research supervisors could not participate directly in these discussions. Final NHSGGC R+D approval was eventually granted on 8<sup>th</sup> September 2009

### **3.4 Participant Screening, Consent and Recruitment Rates**

In order to fully address the proposed hypotheses Dr Dale was required to identify an appropriate cohort of research participants who were willing to participate for the duration of the follow-up period. Potential participants could not be approached, nor consented, until they had undergone a screening review which confirmed their diagnosis and ensured that they fulfilled all of the inclusion / exclusion criteria. The sample size calculation (section 2.2.4) suggested that 110 participants were required to provide sufficient statistical power to address the clinical research hypothesis. It was presumed that approximately 30 percent of patients who met the inclusion / exclusion criteria would prefer not to participate in clinical research and therefore at least 143 patients with early RA would need to be 'screened'. Furthermore, it was imperative to identify

potential participants at the earliest possible time after symptom onset to 1. reduce the risk of their first DMARD being commenced outwith the 'window of therapeutic opportunity' and 2. ensure that as few participants as possible were recruited after the commencement of first DMARD. To maximise recruitment rates, Dr Dale relied on local colleagues kindly agreeing to refer any potential participants who attended their usual new patient and early arthritis clinics. Over the duration of the screening and recruitment period a number of different tactics were employed to try and identify the research cohort as quickly as possible:

*Review of new rheumatology clinic referrals:* All likely new patient referrals to either Gartnavel General Hospital or Stobhill Hospital where the diagnosis of RA was either expressly mentioned, or seemed likely from the provided clinical details, were seen urgently (usually within 2 weeks of referral) by Dr Dale during weekly research clinics at each site or on the rheumatology day-unit at a time convenient to the patient. Patients who fulfilled inclusion / exclusion criteria were offered participation in the research projects; whereas, patients who did not fulfil the inclusion / exclusion criteria were commenced on initial treatment and appropriate local follow-up was arranged

*Referrals from local NHSGGC rheumatologists:* Initially, patients with new diagnoses of early RA, who had attended rheumatology new patient clinics at any of the participating hospitals, were referred to Dr Dale for research screening. Those who met inclusion / exclusion criteria continued to attend the research clinics and on-going follow-up at the local early arthritis clinic was arranged for those who declined. Later, as a means of enhancing recruitment rates, Dr Dale also accepted referrals of potential participants from other local rheumatology services if the patient was willing for their on-going care to be transferred to the closest participating hospital site. In this way, a further 18 research participants were identified from the Victoria Infirmary (5), Inverclyde Royal Hospital (1), Royal Alexandria Hospital (1) and the Private sector (11).

*Attendance at existing NHSGGC early arthritis clinics:* Shortly after commencing recruitment it became apparent that not all patients with new diagnoses of early RA were being referred for screening. Therefore, for a 12 month period, Dr Dale attended the weekly early arthritis clinics at Gartnavel and Glasgow Royal Infirmary as a way of: 1. identifying any patients who might not have been referred previously, 2. assisting local rheumatology colleagues by making the referral pathway as straight forward as possible and 3. raising awareness of the research's aims and inclusion / exclusion criteria

*Recruitment Posters:* A4 posters describing the research's aims, inclusion / exclusion criteria and referral pathways were posted in every clinic room, waiting area and day unit space at each of the participating hospitals

*Email Correspondence:* In order to maintain awareness of the research's aims, inclusion / exclusion criteria and the referral pathways, Dr Dale sent a monthly email reminder to all rheumatology consultants and specialist trainees working in participating and collaborating hospital sites.

The recruitment period was longer than originally anticipated. The first research participant was consented in September 2009 and the final, 111<sup>th</sup> participant was consented in November 2011. Previous, locally administered, studies of early RA, with similar follow-up periods, have recruited similarly sized populations over a similar 3 year period. However, these studies did not face the initial delays in granting of both NHS REC and NHSGGC R+D approvals. Furthermore, the inclusion criteria for these studies allowed patients with symptom durations up to 2 or 5 years to be recruited; whereas, this research limited symptom duration to 1 year thereby excluding a number of patients who might otherwise have helped meet the recruitment target at an earlier point.

### **3.5 Musculoskeletal Ultrasound Training**

Whilst having to wait for final NHS REC and NHSGGC R+D approval did delay the commencement of recruitment it did allow Dr Dale several months to continue practicing and refining the limited MSUS joint set examination which forms the back bone of global disease activity assessment in the MSUS assessment group. Dr Dale received training in MSUS from several different sources:

1. *2008 British Society of Rheumatology Basic Musculoskeletal Ultrasound Training Course.* A three day introductory course delivered by a faculty of British expert rheumatology sonographers
2. *Glasgow Royal Infirmary Rheumatology Department Musculoskeletal Ultrasound Clinic.* Weekly attendance for approximately 9 months for training and supervision on unselected rheumatology out-patients. Provided by Dr Anna Ciechomska, Associate Specialist in Rheumatology, and Dr Debbie Turner, Specialist Podiatrist
3. *2011 Scottish Rheumatology Ultrasound Group Musculoskeletal Ultrasound Training and Teach the Teachers Course.* A three day training course where a faculty of invited, international experts in MSUS delivered training and advice specifically targeted to the needs of the attendees. Part of this course included theoretical and practical synovitis grading exercises which Dr Dale co-organised
4. *Unsupervised Practice.* Over an approximately 12 month period Dr Dale maximised his opportunities to practice and refine the limited MSUS joint set examination by attending the weekly early arthritis clinics at Gartnavel General Hospital, Glasgow Royal Infirmary and Stobhill Hospital as a supernumerary clinician. During these clinics, he would conduct the patient's usual clinic consultation and would complete his assessment of global disease activity by performing the limited MSUS joint set examination. This unsupervised practice proved essential in refining the technique and sequence of joint examination

### **3.6 Identification of Collaborator for Transcriptomic Analysis**

The original intention of the biomarker analysis described in preceding Section 2.7 was to describe changes in expression patterns across a number of different genetic and molecular levels (e.g genotype – RNA / transcriptomic analysis – proteomic – metabolomic). Subsequently, an integrative pathway analysis was proposed, whereby the relative expressions of biomarkers at each level would be compared to the expression levels of corresponding biomarkers on preceding and following levels to determine how whether genetic pre-determinants of RA were transmitted forward to molecular and physical phenotype. This multi-platform analysis requires significant technical assistance from a scientific collaborator with expertise in both operating and conducting each of the different analysis platforms. From the very outset of drafting the research protocol a verbal agreement to collaborate was reached between Dr Dale, his research supervisors at University of Glasgow, and the research scientists of the Translational Medicine Research Collaboration (TMRC), Dundee. The TMRC was funded jointly by Pfizer UK, the Scottish Government and the Universities of Aberdeen, Dundee, Edinburgh and Glasgow. The Research Laboratory provided a state-of-the-art facility with specific expertise in conducting genomic, transcriptomic, proteomic and metabolomic analyses coupled to dedicated bio-informatic support. Unfortunately in January 2010, and despite having virtually finalised a formal research contract and analysis plan, TMRC scientists eventually decided that they would have insufficient capacity to conduct the analysis following the loss of their central funding stream. A potential collaboration with translational immunology scientists based within Pfizer Global's US research laboratories could not progress because of internal conflicts of interest. There then followed a protracted tendering process during which Dr Dale sought quotes from several independent Contract Research Organisations; including, AROS Applied Biotechnology (Aarhus, Denmark), Asuragen Incorporated (Austin, Texas, USA) and Expression Analysis Incorporated (Durham, North Carolina, USA). A potential local collaboration with colleagues at the Sir Henry Wellcome Functional Genomics Unit, University of Glasgow did not progress because of their inexperience at handling human blood samples and the potential scale of the proposed analysis. Eventually, in January 2012, local colleagues (Drs Martin McBride and John McClure) from the Systems Biology Group of the Institute of Cardiovascular and Medical Sciences at the University of Glasgow very kindly agreed to assist by performing the laboratory analysis of the stored blood samples and assisting with the statistical processing. RNA extraction was performed between February and March 2013 and, due to availability of laboratory technical staff, the microarray analysis was completed in August 2012. Data normalisation, quality control procedures and comparison of expression levels between different phenotypic groups was conducted between September 2012 and February 2013.

### **3.7 Recruitment to Scottish Early Rheumatoid Arthritis Inception Cohort**

When the TaSER Study commenced in September 2009 the original intention was that all participants would donate additional blood samples for the proposed multi-level biomarker analysis at set time-points over the duration of their participation in the study. However, whilst the recruitment period of this research was being conducted a further, multi-centre cohort study was commenced with very similar inclusion criteria and sampling requirements. The Scottish Early Rheumatoid Arthritis (SERA) Inception Cohort (UKCRN ID 9162, MREC No 10/S0704/20) is a Scotland-wide, prospective longitudinal study of 1800 patients with early RA, or undifferentiated inflammatory arthritis, which aims to compare changes in clinical and radiographic outcomes to measured changes in biomarker expression profiles to determine whether particular biomarkers (individually or grouped) can act as additional predictors of outcome. Since both projects are analysing outcomes in early RA, to solely recruit patients to the TaSER Study would have significantly impaired recruitment to SERA. Thus, it was agreed that the data collection and blood sampling procedures followed by this research should be brought into line with those of SERA so that participants of the TaSER study might also contribute to the longer term aims of SERA and recruitment to both projects could be maximised. Other than contributing outcome data and blood samples to the SERA study, participants who were primarily recruited to the TaSER study continued to be randomised, assessed and followed-up by the processes described in preceding sections 2.3-2.5. Altogether, of the 111 research participants 79 (study numbers 001-078 and 083) followed the previously described data collection and blood sample donation procedures; 32 participants (study numbers 079-082 and 084-111) contributed clinical and radiological outcome data and blood and urine samples to the SERA study. New study participants could not contribute to the SERA study until the West of Scotland Research Ethics Service had approved a substantial amendment to the research methods. To bring the data collection procedures in line with the main SERA study the following changes were made:

1. Blood biomarker sampling time points switched to baseline and months 6, 12 and 18
2. Additional urine sample collected at each sampling time point for proteomics analysis
3. Completion of the Hospital Anxiety and Depression (HAD) Questionnaire at baseline, 6 and 12 months
4. Additional plain x-rays of hands and feet performed at months 6 and 12
5. Formal documentation of distribution of tender and swollen joints in hands and feet; compared to Ritchie Articular Index component of DAS44 which summarises joint tenderness for the whole metacarpophalangeal and metatarsophalangeal joint areas
6. Documentation of employment status at baseline, 6 and 12 months

## 3.8 Amended Analysis Plan

The previously described delays in obtaining NHS REC and NHSGGC R+D approvals, and in completing the recruitment period, have led to a protracted follow-up period which means that completion of the follow-up period was significantly delayed. Thus, at the time of writing, it has not been possible to report all of the final outcomes described in preceding section 2.6. However, up until this point a significant amount of clinical and ultrasonographic data had been collected; therefore, after discussion with research supervisors and University of Glasgow assessors, a detailed interim analysis was conducted to determine how often the MSUS disease activity findings altered DMARD escalation decisions.

For the purposes of this thesis submission the following results sets will be presented:

*Chapter 4:* Research cohort's presenting demographic and disease-related characteristics. Including: comparison of these characteristics to those of the SERA Study Cohort and other, previously conducted, early RA studies in a similar geographic area

*Chapter 5:* Description of level of agreement between DAS28 and MSUS assessments of global disease activity

*Chapter 6:* Description of the impact of DAS28 and MSUS disease activity assessment on the ACR core set outcomes

*Chapter 7:* Description of findings of initial transcriptomic analysis, conducted using baseline and 3 month PAXgene RNA samples and available clinical data

*Chapter 8:* Description of findings of comparison between baseline, 3 month and 18 month multi-biomarker disease activity test and corresponding DAS28 and MSUS disease activity assessments

### 3.8.1 Interim MSUS Data Analysis Plan

Throughout the follow-up period DAS28 and MSUS disease activity assessment data were frequently recorded simultaneously. These data were interrogated to determine how often DAS28 and MSUS assessment agreed / disagreed on the need for DMARD escalation using the thresholds described in Section 2.4. An estimation of how often MSUS findings were leading to DMARD escalation decisions which differed from those currently reached by standard care (e.g DAS28 assessment) started to demonstrate the likely impact of MSUS assessment upon DMARD treatment intensity and the practicalities of incorporating it into routine practice.

Results from all occasions when there was simultaneous DAS28 and MSUS disease activity assessment data available were pooled. The following relationships were calculated:

1. Proportion of assessment visits where MSUS was indicated and the frequency of each indication. The indications for MSUS were defined as:

- i.  $DAS28 < 3.2$
- ii.  $3.2 \leq DAS28 < 5.1$  and 0-1 swollen joints

Since DMARD escalation is not indicated within 3 months of a preceding DMARD escalation, data for this analysis will be taken from all visits in the MSUS group occurring from month 3 onwards. As an internal comparison, the number of occasions that patients within the DAS28 assessment group might also have qualified for MSUS assessment will be calculated

2. The degree of agreement between DAS28 and MSUS assessments of global disease activity. Agreement will be defined as those occasions when corresponding DAS28 and MSUS assessments lead to the same DMARD escalation decision based on the thresholds described in section 2.4. Disagreement will be defined as those occasions when MSUS findings lead to a DMARD escalation decision that opposes that suggested by the DAS28 findings

- Agreement*
- i.  $DAS28 < 3.2$  and MSUS assessment identifies PD signal in one or no joints. Thus, both assessments agree that DMARD escalation is not indicated
  - ii.  $3.2 \leq DAS28 < 5.1$  and 0-1 clinically swollen joints and MSUS identifies PD signal in two or more joints. Thus, both assessments agree that DMARD escalation is indicated

Overall agreement will be calculated as the proportion of occasions that either scenario i. or scenario ii. are satisfied out of all the occasions when there is corresponding DAS28 and MSUS assessment data available

- Disagreement*
- i.  $DAS28 < 3.2$  but MSUS identifies PD signal in two or more joints. Thus DAS28 suggests low disease activity whereas MSUS assessment has identified in subclinical synovitis
  - ii.  $3.2 \leq DAS28 < 5.1$  and 0-1 clinically swollen joints but MSUS identifies PD signal in one or no joints. Thus, whilst DAS28 suggests moderate disease activity, MSUS has not identified evidence of active synovitis

3. Using MSUS finding data relating to each monthly time point, the proportion of occasions that MSUS findings indicated DMARD escalation will be calculated to determine whether the impact of MSUS alters over the follow-up time period. Furthermore, median synovial hypertrophy and PD signal scores (sum of gradings) and indices (number of positive findings) will be calculated at each time point

### **3.8.2 Amended Transcriptomic Expression Analysis Plan**

So far, it has only been possible to secure local collaborators to perform the transcriptomic analysis, therefore a decision was taken to focus on interrogating the transcriptomic dataset thoroughly before seeking collaborators to assist with analyses across additional polyomic platforms. Analysis was limited to 79 participants out of the whole cohort (study IDs T001 – T078 and T083) since the remaining participants had donated blood samples to the parallel Scottish Early Rheumatoid Arthritis Inception Cohort Study. Furthermore, analysis of transcriptomic expression profiles was restricted to the baseline and 3 month PAXgene RNA samples since, at the time of laboratory analysis, there weren't 18 month samples available for all participants.

Demographic and clinical outcome data - that had been accrued as part of the clinical research - was then used to form patients into pre-defined phenotypic and treatment response groups representing clinical scenarios that rheumatologists were likely to encounter during routine practice. Statistical techniques (described in detail in Section 2.7.2.4) were used to determine whether there were significant differences in the mRNA expression profiles evident between the groups. Comparator groups were formed using the phenotypic descriptors that are described in Section 2.7.2.5

## **4. The TaSER Study Cohort**

## 4.1 Identification of Participants

The recruitment period ran between September 2009 and November 2011 and the follow-up period ran until May 2013. In total, 283 patients were screened and 111 agreed to participate in the research (recruitment rate 39%)

### 4.1.1 Referral Sources

Potential participants could be referred for screening by rheumatology colleagues based at a number of NHSGGC rheumatology units. Participants who were referred from outwith the three main participating sites were allocated appointments at which ever of the main hospital sites was most convenient for their travelling requirements. Table 13 summarises the original source of referral for the final 111 research participants and Table 14 summarises how their follow-up was distributed between each of the participating sites. If, including screening and baseline appointments, all participants had attended for every single monthly review appointment a total of 2,220 consultations would have occurred

Source of Referral	Number of Research Participants	Percentage
Gartnavel General Hospital	49	45%
Glasgow Royal Infirmary	24	21%
Stobhill Hospital	20	18%
Private Practice	11	10%
Victoria Infirmary	5	4%
Inverclyde Royal Hospital	1	<1%
Royal Alexandria Hospital, Paisley	1	<1%

**Table 13:** Original source of referral for research participants

Follow-up Site	Number of Research Participants	Percentage
Gartnavel General Hospital	63	57%
Glasgow Royal Infirmary	26	23%
Stobhill Hospital	22	20%

**Table 14:** Follow-up arrangements for research participants

### 4.1.2 Screening Process

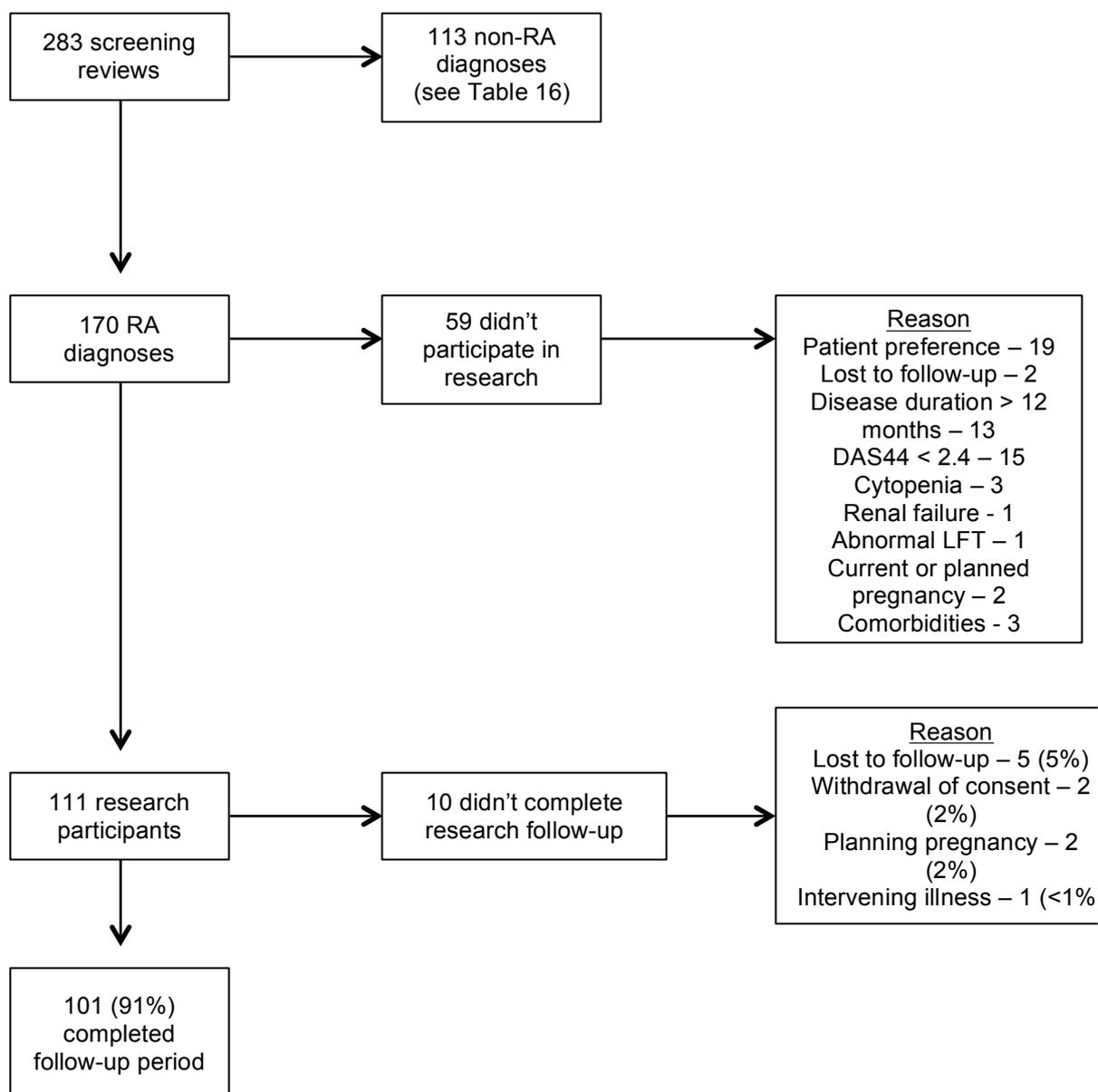
283 patients underwent screening review; of these 170 (60%) had a clinical diagnosis of RA and 113 (40%) had an alternative diagnosis. Of the RA patients, 111 (65%) agreed to participate in the research (overall recruitment rate = 39%) and 59 (35%) did not, or could not, participate. Table 15 summarises the reasons why RA patients were unable to participate in the research. Table 16 describes the range of alternative diagnoses of non-RA patients who attended screening. The consort diagram in Figure 12 depicts how patients and research participants have progressed through the screening and follow-up process to date

Reason	Frequency	Percentage
Patient preference	19	32%
Did not attend follow-up	2	3%
<i>Did not meet inclusion criteria</i>		
Disease duration > 12 months	13 (mean = 30 months)	22%
DAS44 < 2.4 (after steroid therapy)	15 (5)	25% (8%)
<i>Fulfilled exclusion criteria</i>		
Cytopenia (Hb < 10g/dl, WCC < 4x10 <sup>9</sup> /l, platelet < 150x10 <sup>9</sup> /l)	3	5%
Renal failure (creatinine > 200 µg/l)	1	2%
Abnormal liver function tests (AST / ALT > twice ULN)	1	2%
Pregnant or planning pregnancy	2	3%
Aggressive DMARD escalation contraindicated by co- morbidities	3	5%

**Table 15:** Frequency of reasons why patients with RA (n=59) did not participate in research

Diagnosis	Frequency	Percentage
Osteoarthritis	22	19%
Undifferentiated inflammatory arthritis	17	15%
aCCP / RhF +ve – no arthritis	16	14%
Psoriatic arthritis	9	8%
Fibromyalgia	9	8%
Arthralgia (no specific diagnosis)	6	5%
Palindromic arthritis	5	4%
Reactive arthritis	3	3%
Transient arthritis	3	3%
Ankylosing spondylitis	2	2%
Gout	2	2%
Sarcoidosis	2	2%
Connective tissue disease	2	2%
Viral arthritis	2	2%
Extensor tenosynovitis	2	2%
Sjogren's syndrome	2	2%
Hypermobility syndrome	1	<1%
CTD associated polymyositis	1	<1%
Ulnar nerve palsy	1	<1%
Flexor tenosynovitis	1	<1%
Diabetic cheiroarthropathy	1	<1%
IBD related arthritis	1	<1%
Polymyalgia rheumatic	1	<1%
Meniscal tear	1	<1%
Dupuytren's contracture	1	<1%

**Table 16:** Frequency of alternative diagnoses in patients (n = 113) attending for screening review



**Figure 12:** Consort diagram depicting progress of patients and research participants through screening and follow-up period

Ten (9%) patients withdrew before completing the full follow-up period: 5 (5%) had not attended for follow-up reviews; 2 (2%) had withdrawn consent and elected not to continue in the research and 2 (2%) were planning to conceive. Additionally, 1 patient developed an overlap syndrome with dermatomyositis and required such high levels of oral corticosteroids (initially prednisolone 60mg/d) and alternative immunosuppressant therapy (azathioprine then rituximab) that it was not possible to continue following the DMARD escalation protocol

## 4.2 Baseline Characteristics of Study Cohort

Overall, the research cohort appeared representative of a typical early RA population. Table 17 describes the baseline demographic and disease related characteristics of the whole cohort in detail. The cohort can be broadly summarised as predominantly female (68%) and middle aged (mean age 56 years) with recent onset symptoms (mean symptom duration 5.3 months); moderate-to-high disease activity (mean baseline DAS44 = 4.3), moderate functional impairment

(mean baseline HAQ = 1.5). The majority exhibited either positive rheumatoid factor and/or anti-CCP antibodies (67% and 60% respectively). Despite the relatively early presentation a significant minority (29%) still exhibited erosive baseline x-rays.

	TaSER			p (DAS28 vs US)
	All	DAS28	Ultrasound	
<b>Number</b>	111	57	54	
<b>Females</b> n (%)	76 (68%)	43 (75%)	33 (61%)	0.10 (chi <sup>2</sup> )
<b>Age</b> (years)	56 (13)	56 (13)	57 (14)	0.68
<b>Disease Duration</b> (months)	5.3 (3.0)	5.4 (3.1)	5.1 (2.8)	0.64
<b>RhF +ve</b> n (%)	74 (67%)	39 (68%)	35 (65%)	0.16 (chi <sup>2</sup> )
<b>CCP +ve</b> n (%)	67 (60%)	35 (61%)	32 (59%)	0.81 (chi <sup>2</sup> )
<b>Current Smoker</b> n (%)	31 (28%)	17 (30%)	14 (26%)	0.728
<b>Ex-smoker</b> n (%)	31 (28%)	18 (32%)	13 (24%)	
<b>Never Smoked</b> n (%)	49 (44%)	22 (39%)	27 (50%)	
<b>DAS28</b>	5.0 (1.1)	5.0 (1.2)	4.9 (1.0)	0.56
<b>28TJC</b>	6 (5)	6 (5)	6 (5)	0.85
<b>28SJC</b>	6 (4)	6 (4)	6 (3)	0.68
<b>DAS44</b>	4.3 (1.2)	4.4 (1.3)	4.3 (1.0)	0.59
<b>RAI</b>	20 (13)	20 (14)	20 (12)	0.92
<b>44SJC</b>	10 (6)	10 (7)	9 (5)	0.38
<b>Patient Global</b> (100mm VAS)	54.7 (22.3)	56.2 (27.9)	54.4 (21.4)	0.71
<b>Physician Global</b> (Likert 0-5)	3.2 (0.7)	3.2 (0.7)	3.2 (0.7)	0.65
<b>ESR</b> (mm/hr)	39 (35)	41 (43)	36 (25)	0.41
<b>CRP</b> (mg/l)	37 (49)	38 (57)	35 (40)	0.76
<b>HAQ</b>	1.5 (0.8)	1.5 (0.8)	1.5 (0.7)	0.92
<b>Pain VAS</b> (100mm VAS)	49.4 (22.4)	52.5 (25.1)	46.2 (18.9)	0.14
<b>Baseline Erosions*</b> n (%)	32 (29%)	15 (26%)	17 (32%)	0.470

**Table 17:** Baseline demographic and disease related characteristics of whole research cohort, DAS28 and MSUS assessment groups

Unless otherwise stated values are shown as means ± standard deviation

\*Data collected from routine NHSGGC radiology reports, formal grading of baseline xrays is awaited

Following the randomisation process, the DAS28 group was slightly larger than the MSUS group and contained a numerically higher proportion of female participants (75% vs 61%, p=0.10). Other than gender, there were no statistically significant between group differences evident in any of the baseline demographic features, disease characteristics or measures of disease activity.

#### *Autoantibody Status*

Observational studies suggest that the prevalence of anti-CCP positivity amongst European early RA cohorts is approximately 57 – 63% and for rheumatoid factors is approximately 54 – 63% (1,10); reassuringly, very similar results were observed in this cohort (60% anti-CCP positive, 67% RhF positive). Table 18 describes the rates of separate and combined autoantibody positivity for all research participants. These rates are comparable to those previously described in 164 RA patients attending routine early arthritis clinics in Glasgow (65% anti-CCP positive, 63% RhF positive) (J Dale – unpublished data). Additionally, the rates of combined anti-CCP antibody and/or rheumatoid factor status for this cohort were very similar to those described in a larger (n=279) Swedish early RA cohort who presented with similar symptom durations (mean = 5 months) and disease activity levels (mean DAS28 = 5.01) (1).

All patients		Anti-CCP antibody		Total
		Negative	Positive	
Rheumatoid Factor	Negative	32 (29%) (32%)	5 (4%) (4%)	37 (33%)
	Positive	12 (11%) (10%)	62 (56%) (53%)	74 (67%)
	Total	44 (40%)	67 (60%)	111

**Table 18:** Frequency of anti-CCP and/or rheumatoid factor positivity for the whole research cohort

Values are number (percentage).

Figures in red font show corresponding rates for the Swedish RA cohort (1)

### **4.2.1 Fulfilment of RA classification criteria**

#### *1987 ACR RA Classification Criteria*

Since the 1987 ACR classification criteria for RA are biased towards identifying established RA, eligible participants needed a robust clinical diagnosis of RA (based on presenting features and initial investigation results) rather than needing to fulfil the criteria set. Even so, data relating to whether or not participants fulfilled 1987 ACR classification criteria was still collected during each baseline assessment. Since some participants presented very early after the onset of symptoms, with clear features of RA, there was no requirement for criteria 1-4 to have been present before the criteria could be applied. Based on presenting features, 95 participants (86%) fulfilled 1987 ACR classification criteria for RA (i.e. total score  $\geq 4$ ). Table 19 describes the frequency that each criterion was present and Table 20 describes how often each score was returned. Of note, the criteria which are most closely related to disease longevity (e.g. presence of rheumatoid nodules, erosive baseline x-rays) were recorded much less frequently (4% and 29% respectively) than the other criteria which relate to presenting clinical features and fixed rheumatoid factor status

Criterion	Frequency	Percentage
Early morning stiffness $\geq$ 1 hour	94	85%
Arthritis $\geq$ 3 areas	102	92%
Arthritis of hand joints	105	95%
Symmetric arthritis	95	86%
Presence of rheumatoid nodules	4	4%
Rheumatoid factor positive	75	67%
Erosive plain x-rays	32	29%

**Table 19:** Fulfilment of 1987 ACR Classification Criteria – during baseline assessment  
Definitions of criteria adapted from Arnett et al (11). The '6 week rule' for criteria 1-4 was not applied

Score	Frequency	Percentage
2	1	1%
3	15	14%
4	34	31%
5	44	40%
6	15	14%
7	2	2%

**Table 20:** Fulfilment of 1987 ACR Classification Criteria – total score at baseline assessment  
A score  $\geq$  4 classifies the presenting features as RA(11)

#### *2010 ACR-EULAR Classification Criteria for Rheumatoid Arthritis*

During the recruitment process the 2010 ACR-EULAR Classification Criteria for RA were proposed (13). The 2010 classification criteria were not specifically included within the original dataset; however, fortuitously each of the individual components was already being collected during the baseline assessment and it was therefore possible to retrospectively apply the 2010 classification criteria. The distribution of large and small joint involvement was determined from the 44 swollen joint count and Ritchie Articular Index components of the DAS44. The Ritchie Articular Index summarises tenderness across all PIPj and MCPj of each hands and therefore doesn't provide sufficient detail to accurately determine the extent of small joint involvement in the hands. In instances where the Ritchie Articular Index identified PIPj or MCPj tenderness the extent and pattern of joint involvement was determined by referring to the 28 tender joint count that was recorded independently as part of the baseline DAS28 assessment. All study participants displayed evidence of either clinical and/or musculoskeletal ultrasound synovitis in at least one joint at the baseline assessment.

One hundred and seven participants (96%) scored 6 or higher and therefore satisfied 2010 ACR-EULAR Classification Criteria for definite RA. Table 21 describes the frequency that each criterion was present and Table 22 demonstrates the frequency that each score was reported

Criterion	Score	Frequency	Percentage
<b>A. Joint Involvement</b>			
1 large joint	0	8	7%
2-10 large joints	1	24	22%
1-3 small joints*	2	12	11%
4-10 small joints*	3	24	22%
>10 joints (at least 1 small joint)	5	74	67%
<b>B. Serology</b>			
Negative RF and negative ACPA	0	32	29%
Low positive RF or low positive ACPA	2	7	6%
High positive RF or high positive ACPA	3	72	65%
<b>C. Acute-phase Reactants</b>			
Normal CRP and normal ESR	0	19	17%
Abnormal CRP or abnormal ESR	1	92	83%
<b>D. Duration of Symptoms</b>			
< 6 weeks	0	5	5%
≥ 6 weeks	1	106	95%

**Table 21:** Fulfilment of 2010 ACR-EULAR Classification Criteria for RA – during baseline assessment.

\* with or without large joint involvement

Definitions of criterions based upon Aletaha et al (13)

Score	Frequency	Percentage
5	4	4%
6	11	10%
7	28	25%
8	10	9%
9	24	22%
10	34	30%

**Table 22:** Fulfilment of 2010 ACR-EULAR Classification Criteria for RA – total score at baseline assessment.

A total score ≥ 6 identifies definite RA

Altogether, 95 participants (86%) fulfilled the 1987 ACR criteria, 107 participants (96%) fulfilled the 2010 ACR-EULAR criteria and 92 participants (83%) fulfilled both sets of criteria. Three participants (3%) solely fulfilled the 1987 ACR criteria, 15 participants (14%) solely fulfilled the 2010 ACR-EULAR criteria and 1 participant (approx. 1%) fulfilled neither set of criteria. Chi squared comparison showed no significant statistical difference between the proportion of participants in each randomisation group fulfilling either set of criteria (p=0.54).

Taken together these results provide additional support that the screening process identified an appropriate study cohort to follow since the vast majority of participants satisfied current RA classification criteria for RA. Furthermore, the overall picture implied by the individual criterion does fit with a very typical description of RA. The majority of participants displayed 1. polyarticular involvement at presentation (67% with >10 clinically affected joints); 2. high titres of at least one disease associated autoantibody (65% high titre RhF and/or CCP); 3. elevated acute phase reactants (83% abnormal CRP and/or ESR) and 4. persistent disease outwith the expected window for spontaneous resolution (95% > 6 weeks symptom duration).

## 4.2.2 Distribution of Joint Involvement at Presentation

RA is traditionally considered to involve peripheral small joints in a symmetrical distribution (4). However, this definition probably relates better to established disease since it is increasingly recognised that the earliest stages of symptoms are not always associated with symmetrical joint involvement. Indeed, the 2010 ACR-EULAR RA classification criteria no longer feature symmetrical joint involvement as a defining feature (13) and recent prediction rules for persistent, erosive arthritis allocate either no, or a relatively small, additional weighting to the presence of symmetrical joint involvement (15,16). To determine how closely participant's presenting features matched the classical description of RA the distribution of clinically evident synovitis in peripheral joints was recorded during the baseline assessment (Table 23). Interestingly, despite the relatively short symptom durations, the distribution of joints involved did still suggest that significant numbers of the participants presented with clinical features that fitted the classical description of rheumatoid arthritis (e.g, symmetrical involvement of peripheral small hand and foot joints). The most common patterns of clinical joint inflammation were bilateral involvement of MCPj (70%), PIPj (64%), MTPj (53%) and wrist (49%).

		No Involvement		Unilateral Involvement		Bilateral Involvement	
		Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Upper Limb	PIPj	29	26%	11	10%	71	64%
	MCPj	25	23%	8	7%	78	70%
	Wrist	34	31%	23	21%	54	49%
	Elbow	92	83%	10	9%	9	8%
	Shoulder	65	77%	16	14%	30	27%
Lower Limb	MTPj	44	40%	8	7%	59	53%
	Ankle	85	77%	8	7%	18	16%
	Knee	61	55%	28	25%	22	20%
	Hip	109	98%	0	0	2	2%

**Table 23:** Frequency and distribution of clinically evident synovitis at presentation

## 4.3 Discussion

Before the impact of the intervention can be assessed the ability of the research cohort to represent a 'typical' early RA population should be considered. Newly diagnosed RA patients who were undergoing screening were primarily considered suitable for recruitment if their presenting clinical and laboratory features reasonably supported a clinical diagnosis of RA. Whilst this approach reduces any potential restrictions to recruitment that might be caused by structured classification criteria it does increase the risk of diagnostic inconsistency because of clinician's subjectivity.

### 4.3.1 Baseline Clinical Features

Overall, the baseline clinical and laboratory descriptors of the TaSER cohort seem typical of a newly diagnosed early RA population. At a whole cohort level, the participants presented in the sixth decade (mean age 56 years) and the majority (68%) were female. Participants presented relatively early (mean = 5.3 months) after the onset of symptoms and whilst the majority of presentations were from a fairly narrow central band (IQR 3 – 7 months) the overall range was much wider (0.5 – 12 months). Participants presented with moderate-high clinical disease activity

(mean DAS44 4.4), moderate elevation of inflammatory markers (mean ESR 39, mean CRP 37), moderate functional impairment (mean HAQ 1.5) and elevated participant completed assessments of pain and global health (mean 10cm VAS 49/100 and 55/100 respectively). The majority of participants returned HAQ assessments (IQR 1-2; SD 0.77) pain 10cm VAS assessments (IQR 34-63; SD 22.4) and global health 10cm VAS assessments (IQR 38.5 -71; SD 22.3) from within a narrow central range; however, the total ranges for each of these variables (HAQ 0-2.9; pain VAS 8-100; global health VAS 7-100) do suggest that there were small numbers of outlying participants at both extremes of each variable.

Excepting sex, there were no statistical differences in the baseline measures between the assessment groups. The randomisation process led to the DAS28 assessment group being slightly larger (57 vs 54) than the MSUS assessment group and containing a numerically higher proportion of females (75% vs 61%). It is quite likely that this imbalance has resulted from the minimisation processes used to ensure equal balance of DAS28, rheumatoid factor status and erosive status between the assessment groups. Being female does have clear prognostic implications (32,320) and several population based studies have demonstrated that measures of disease activity and functional impairment tend to be worse in females (36,38-40). It is therefore possible that the DAS28 group may have been predisposed to achieving a lesser treatment response and poorer outcomes over the course of the follow-up period. Nevertheless, all other baseline prognostic markers (age, rheumatoid factor and anti-CCP antibody status, symptom duration at presentation, smoking status, baseline disease activity level) were equally represented in both assessment groups and may therefore limit any potential bias associated with the imbalance in female sex. In fact, an increased proportion of females within the DAS28 group should have increased the likelihood of identifying between group differences if the control group were effectively biased towards worse outcomes. Analyses performed to date have shown that there were no differences in DAS44 or gene expression outcomes between the gender groups, though male participants did tend to demonstrate earlier and greater improvements in functional ability. In due course, gender will be incorporated into multiple regression analyses of the clinical outcomes to determine whether it comes out as an independent predictor of clinical response.

Analysis of the pattern of clinical joint involvement at presentation suggests that the cohort's presentation was consistent with RA. There was polyarticular joint involvement (mean 44 swollen joint count 10) and the small peripheral joints of the hands and feet were most commonly involved (any involvement: MCP 77%, hand PiPj 74%; wrist 69%; MTPj 60%). Furthermore, during the baseline assessment there was consistently high rates of bilateral involvement for all joint areas, except the elbows and knees (Table 23). There is increasing recognition that patients in the very early stages of RA, when the disease process may be most amenable to treatment, do not necessarily present with symmetrical joint involvement. Indeed, a recent cohort study conducted on 2472 patients with early undifferentiated arthritis suggested that those who fulfilled the 2010 ACR-EULAR RA classification criteria presented with earlier disease (mean symptom duration: 13.4 vs 15.7 weeks;  $p = 0.01$ ), were more likely to exhibit asymmetrical joint involvement (52% vs 46%;  $p = 0.01$ ) and had a higher rate of small hand and foot joint involvement (85% vs 81%;  $p = 0.04$ ) than those who fulfilled the 1987 ACR classification criteria (321). Few recent studies have

described the distribution of joint involvement in newly diagnosed RA patients or those with early undifferentiated arthritis. Compared to the cohort study, the TaSER cohort exhibited a somewhat higher rate of symmetrical joint involvement which could, in part, have been related to the inclusion criteria requiring potential participants to have a clinical diagnosis of RA. It is possible that this approach emphasised the need for potential participants to match the 'classical' description of RA and thus predisposed towards recruiting participants with symmetrical joint involvement.

### **4.3.2 Baseline Autoantibody Status**

It is now well recognised that rheumatoid factor and (especially) anti-citrullinated protein antibodies (ACPA) have important diagnostic and prognostic properties in relation to RA (6,10,69). Therefore, the expression of these autoantibodies within an early RA research cohort should be comparable to both the local RA population and other previously described early RA cohorts. It is also understood that rheumatoid factor and/or anti-CCP antibodies are detectable for several years before the development of RA and that the prevalence of either autoantibody does vary depending upon which stage of RA is present. Prospective observational studies have demonstrated that i. the prevalence of rheumatoid factor and anti-CCP antibodies in the asymptomatic phase prior to the development of clinical RA is between 19.3 – 57% and 33.7 – 61% respectively (7,84,85) and that ii. the prevalence of both antibodies increases sharply immediately before the onset of clinical RA (7). Similarly, in newly diagnosed RA the prevalence of rheumatoid factors and anti-CCP antibodies is significantly higher than in asymptomatic at-risk patients (84) and, furthermore, is higher still in patients with longer disease durations and well established diagnoses of RA (IgM RF 68-82%; anti-CCP 72-82%) (322-324).

For the TaSER cohort the prevalence of rheumatoid factor was 67% and anti-CCP2 antibodies was 60%. Rheumatoid factor and anti-CCP status was concordant in 94 (85%) participants. These results are at least comparable to the baseline values reported in local NHSGGC early arthritis clinics and similar early RA cohorts.

- In 2010, a retrospective analysis was conducted of the case sheets of 164 RA patients attending NHSGGC early arthritis clinics to determine their rate of rheumatoid factor and anti-CCP antibody positivity (J Dale – unpublished results). Table 24 summarises the results of this analysis. Overall, the rates of anti-CCP antibody (60% vs 63%) and paired RhF-anti-CCP positivity (56% vs 54%) were similar between the TaSER cohort and NHSGGC clinics. The NHSGGC clinics contained a notably higher proportion of RhF positive patients (67% vs 80%) which may reflect false positive results causing diagnostic bias since, at the time of analysis, none of the clinics were applying RA classification criteria to confirm the clinical diagnosis

		<b>Anti-CCP Positive</b>	<b>Anti-CCP negative</b>	<b>Total</b>
<b>TaSER Cohort</b> N = 111	<b>RhF Positive</b>	62 (56%)	12 (11%)	74 (67%)
<b>NHSGGC Cohort*</b> N = 164		89 (54%)	43 (26%)	132 (80%)
<b>TaSER Cohort</b>	<b>RhF Negative</b>	5 (4%)	32 (29%)	37 (33%)
<b>NHSGGC Cohort</b>		15 (9%)	17 (10%)	32 (19%)

**Table 24** - Comparison of rates of rheumatoid factor and anti-CCP antibody positivity between TaSER study cohort and NHSGGC early arthritis cohort

- The Scottish Early Rheumatoid Arthritis Inception Cohort Study (SERA – UKCRN ID 9162) provides a useful parallel cohort, from a similar geographic area, against which the baseline features of the TaSER cohort can be compared. The following results are a personal communication from Dr D Porter (SERA Principal Investigator) and are presently unpublished. 508 (78%) SERA participants had been classified as RA by the 2010 ACR-EULAR RA classification criteria. Overall, the incidence of RhF and/or anti-CCP antibodies appears similar between the TaSER and SERA cohorts. Rheumatoid factor results were available for 367 patients (57% of the total cohort), of whom, 322 patients (88%) are rheumatoid factor positive. Anti-CCP antibody results are available for 525 patients (81% of the total cohort), of whom, 346 patients (66%) are anti-CCP antibody positive. At present it is not possible to report the degree of concordance between rheumatoid factor and anti-CCP antibody.
- Several recent cohort studies have reported the incidence of RhF and anti-CCP antibody positivity in early and established RA populations from different geographical regions (1,323,324). Overall, rates of RhF and anti-CCP antibody positivity and concordance were relatively similar between the TaSER study and other cohorts. There was evidence of some geographical variation with RA patients from India (324) and Brazil (323) tending to have slightly higher rates of dual positive RhF and anti-CCP antibodies whilst patients from Sweden (i.e. another North European cohort) had very similar rates of single and dual antibody positivity (1) (Table 25). Of note, results reported from the Indian cohort were gathered from patients with established RA where the incidence of RhF and anti-CCP antibodies may be expected to be higher anyway.

		<b>Anti-CCP Positive</b>	<b>Anti-CCP negative</b>	<b>Total</b>
<b>TaSER Cohort</b> N = 111	<b>RhF Positive</b>	62 (56%)	12 (11%)	74 (67%)
<b>Swedish Cohort*</b> N = 279		147 (53%)	28 (10%)	175 (63%)
<b>Indian Cohort**</b> N = 129		95 (74%)	11 (9%)	106 (82%)
<b>Brazilian Cohort***</b> N = 156		95 (61%)	11 (7%)	106 (68%)
<b>RhF Negative</b>				
<b>TaSER Cohort</b>	<b>RhF Negative</b>	5 (4%)	32 (29%)	37 (33%)
<b>Swedish Cohort</b>		12 (4%)	90 (32%)	102 (36%)
<b>Indian Cohort</b>		11 (8%)	12 (9%)	23 (18%)
<b>Brazilian Cohort</b>		22 (14%)	18 (12%)	40 (26%)

**Table 25** – Comparison of rates of rheumatoid factor and anti-CCP antibody positivity between TaSER study cohort and Swedish early arthritis cohort

\* Adapted from Ronnelid et al. Annals of Rheumatic Diseases 2005 (1)

\*\* Adapted from Agrawal et al. Clinical Rheumatology 2007 (324)

\*\*\* Adapted from Goeldner et al. Rheumatology 2010 (323)

Recent evidence suggests that a patient's anti-citrullinated protein antibody (ACPA) status is dynamic with the pattern of ACPA expression undergoing a two-fold change prior to the onset of clinical RA. Observational studies of ACPA-positive individuals without RA have demonstrated that, an individual's total ACPA titre rises prior to the onset of RA(7) and also diversifies with expression of antibody isotypes directed against a wider range of citrillunated antigens (epitope spreading)(87,88). Whilst the dynamic changes in ACPA expression are extremely interesting it is not yet possible to discuss how ACPA isotype expression may have influenced either the presenting features or observed outcomes of the TaSER cohort. The anti-CCP assay used by NHSGGC Immunology department is a second generation test which only quantifies total anti-CCP expression and not the expression of individual ACPA isotypes. Recently, a verbal agreement to collaborate has been reached with Dr Jeremy Sokolove (Division of Rheumatology, Stanford University Medical School) whereby baseline serum samples will be analysed using the Stanford custom multiplex autoantibody and cytokine array. Correlation and regression analyses will then be conducted to determine how strongly baseline autoantibody profile predicts subsequent clinical course and treatment response.

### 4.3.3 Fulfillment of Existing RA Classification Criteria

To determine how closely the research cohort matched existing definitions of RA the 1987 ACR RA classification criteria (11) and the 2010 ACR-EULAR RA classification criteria (13) were applied to the baseline presenting features of the TaSER cohort. Reassuringly, both the 1987 ACR criteria and the 2010 ACR-EULAR criteria classified a high proportion of patients as RA (86% and 96% respectively). Furthermore, examining the individual criterions of each criteria set once again suggested baseline features that were compatible with the common perception of an RA phenotype. From the 1987 ACR criteria, the majority of patients described prolonged morning stiffness (>1 our) and exhibited a symmetrical arthritis (86%), frequently involving the hand joints

(95%) and often associated with rheumatoid factor positivity (67%) (Table 19). Markers of established disease, such as rheumatoid nodules (4%) and erosive baseline xrays (29%), were not prevalent. A similar impression is created by the 2010 ACR-EULAR classification criteria: the majority of participants exhibited polyarticular joint involvement (>10 joints 67%), high RhF and or anti-CCP antibodies (65%) and elevated acute phase reactants (83%).

Most recent early arthritis intervention studies require patients to fulfil existing RA classification criteria (usually 2010 ACR-EULAR criteria) before they can be considered for recruitment. However, in validation studies the gold standard against which the performance of the classification criteria is compared is often either subjective (e.g. physician's opinion) or non-specific (e.g. need to commence DMARD therapy). Several studies have reproduced a similar situation to the screening and recruitment process of this research; whereby, the performance of the 1987 and 2010 classification criteria are compared to physician's diagnosis of RA (325-327). Indeed, when both classification criteria are applied strictly to early inflammatory arthritis cohorts the 2010 ACR-EULAR criteria are consistently found to be more sensitive at identifying 'definite RA.' In particular, compared to the 1987 criteria, the 2010 ACR-EULAR criteria are more likely to classify patients with asymmetric oligoarthritis presentations as RA (325) and are more sensitive in patients who are seronegative for RA associated antibodies (328). Furthermore, a recent meta-analysis of 5 classification studies has suggested that whilst the 2010 ACR-EULAR criteria are more sensitive for the diagnosis of RA, the 1987 ACR criteria appear to be more specific (327)

For this study, very high proportions of the cohort fulfilled both the 1987 ACR and 2010 ACR-EULAR classification criteria for RA. This provides additional supporting evidence that the screening and recruitment process identified an appropriate cohort of patients. In line with previous studies, a slightly higher proportion of participants fulfilled the 2010 ACR-EULAR criteria than the 1987 ACR criteria; however, this did not reach statistical significance. To ensure that only patients with early RA, rather than any cause of inflammatory arthritis were recruited, the inclusion criteria stipulated a clear clinical diagnosis of RA, or anti-CCP positive undifferentiated inflammatory arthritis. Therefore, it is unsurprising that a high proportion of participants fulfilled one or both sets of classification criteria, since the referring clinicians and study investigators will have been biased towards considering patients where the clinical diagnosis was apparent and in keeping with the 'classical' understanding of RA. Indeed, it is in patients with undifferentiated inflammatory arthritis where the 2010 ACR-EULAR classification criteria are most likely to be clinically useful. A high proportion of participants fulfilling the 1987 ACR criteria may suggest that a significant number of participants were recruited at a later stage of the disease process. In fact, the mean symptom duration for the whole cohort was relatively short (5.3 months) and was comparable (5.2 months) when considering the subset of 95 patients who fulfilled the 1987 criteria.

#### **4.3.4 Characteristics of other recent early RA treatment strategy studies**

To determine whether MSUS-driven DMARD escalation has a similar impact upon outcomes to other interventional strategies the presenting features of the TaSER cohort were compared numerically to those of patients recruited to other strategic treatment studies (Table 26). The TICORA (118) and TEAR (UK) (146) studies were conducted in exactly the same geographical area though had notable differences of inclusion criteria. Both studies recruited participants with symptom durations up to 5 years, whilst the TEAR (UK) study required participants to have at least high disease activity (DAS28 > 5.1). The SWEFOT study (147) inclusion criteria also allowed symptom durations up to 1 year but did not specify a lower disease activity limit. Overall, allowing for differences in disease duration, the baseline disease characteristics and disease activity measures of the TaSER cohort appear numerically similar to previous early interventional cohorts

Taken altogether, when the baseline features of patients recruited to the TaSER are examined closely it is possible to be confident that a representative population of early RA patients was identified by the screening and recruitment process. The baseline clinical features, demographic and disease related characteristics of the cohort is similar to well-established definitions of the presenting features of RA and very high proportions of participants fulfilled either one or both of the accepted classification criteria definition for RA. Excepting female sex, there were no significant differences in the distribution of baseline demographic or disease-related phenotypic variables between the DAS28 and MSUS assessment groups suggesting that the baseline features of either group are unlikely to negatively bias the measured impact of the assessment strategy. Allowing for heterogeneity of study design (in particular, variations in inclusion criteria), and without being able to conduct direct statistical comparisons, comparisons with other published studies do suggest that the baseline RA-related phenotypic variables for this research cohort are at least numerically similar to those of other early RA cohorts. Rates of RA-associated auto-antibody expression were similar to those of other locally identified and European early RA cohorts. Furthermore, excepting the TEAR (UK) study, this research cohort exhibited similar baseline demographic features and disease activity levels to those of several, previously published, early RA inception cohorts.

The screening process also identified 59 additional RA patients who, through preference or presence of a contra-indication, were unable to participate in the research. If recruited these patients would have comprised a significant proportion of the overall cohort and their individual disease characteristics could feasibly have altered the baseline characteristics and treatment response of both the whole cohort and the individual assessment groups.

	TaSER		TICORA <sup>94</sup>		TEAR (UK) <sup>231</sup>		SWEFOT <sup>112</sup>	
	Region	Time period	Region	Time period	Region	Time period	Region	Time period
	Glasgow, Scotland	Glasgow, Scotland	Glasgow, Scotland	Glasgow, Scotland	Glasgow, Scotland	Glasgow, Scotland	Sweden	
	September 2009 – May 2013	September 2009 – May 2013	August 1999 – April 2001	August 1999 – April 2001	February 2003 – March 2005	February 2003 – March 2005	October 2002 – December 2005	October 2002 – December 2005
	Ultrasound	DAS28	Intensive	Routine	Step-Up	Parallel	DMARD	Infliximab
<b>Number</b>	53	58	55	55	47	49	130	128
<b>Females (%)</b>	60%	78%	71%	69%	79%	76%	78%	76%
<b>Age</b>	57 (13)	55 (14)	51 (15)	54 (11)	55 (11)	55 (15)	53 (14)	51 (13)
<b>Symptom Duration (months)</b>	5.1 (2.8)	5.4 (3.1)	19 (16)	20 (16)	13 (12)	10 (9)	6.3 (3.6)	6.2 (3.5)
<b>RhF +ve</b>	66%	67%	75%	73%	72%	69%	65%	69%
<b>CCP +ve</b>	66%	60%						
<b>DAS28</b>	5.0 (1.1)	5.0 (1.2)			6.9 (0.9)	6.8 (0.9)	5.98 (0.96)	5.91 (0.93)
<b>DAS44</b>	4.4 (1.0)	4.3 (1.3)	4.9 (0.9)	4.6 (1.0)				
<b>Patient Global 100mm VAS</b>	53 (21)	57 (24)	69 (21)	62 (23)	77 (18)	82 (18)		
<b>Physician Global</b>	3.2 (0.7)	3.2 (0.7)	70 (18)	65 (18)	3.6 (0.6)	3.4 (0.6)		
<b>ESR (mm/hr)</b>	37 (26)	40 (42)	45 (31)	34 (27)	45 (30)	36 (24)		
<b>CRP (mg/l)</b>	37 (41)	37 (56)	44 (53)	38 (50)	58 (53)	38 (41)		
<b>HAQ</b>	1.6 (0.7)	1.5 (0.8)	2.0 (0.8)	1.9 (0.7)	2.0 (0.7)	1.9 (0.7)	1.32 (0.60)	1.27 (0.60)
<b>Pain VAS (100mm VAS)</b>	46.8 (20.0)	51.8 (24.6)	62 (20)	59 (20)	65 (22)	71 (26)		

**Table 26** - Baseline RA disease characteristics and disease activity levels for patients recruited to the TASER, TICORA, TEAR (UK) and SWEFOT studies of RA treatment strategies

# **5. Impact of Musculoskeletal Ultrasound on DMARD Escalation Decisions**

Depending upon their clinical course participants within the MSUS group underwent MSUS assessment of global disease activity on a highly variable number of occasions. The following sections will describe the degree of agreement between DAS28 and MSUS assessments of global disease activity and the potential impact of using the proposed indications and definitions for MSUS assessment on both clinical workload and treatment escalation decisions. In situations where MSUS and DAS28 assessment disagreed about the disease activity state, MSUS findings were given precedence over DAS28. Hence, it should be assumed that DMARD therapy would have been escalated if the MSUS escalation threshold was exceeded and provided there was no specific contraindication identified. Follow-up months 1 and 2 are excluded from the analyses since participants would have been within 3 months of commencing DMARD therapy and therefore would not have been eligible for consideration of DMARD escalation.

## 5.1 Fulfilment of Indication Criteria for MSUS

Table 27 summarises the different disease activity states that were used to categorise participants' DAS28 findings and how these related to DMARD escalation decisions

Disease Activity State Descriptor	Disease State	DMARD escalation decision
DAS28 $\geq$ 5.1	High disease activity	Escalate on clinical grounds
$3.2 \leq$ DAS28 < 5.1 AND 28SJC $\geq$ 2	Moderate disease activity and clinical synovitis	
$3.2 \leq$ DAS28 < 5.1 AND 28SJC = 0-1	Moderate disease activity but minimal clinical synovitis	MSUS assessment required to inform DMARD decision
DAS28 < 3.2	Low disease activity	
DAS28 < 2.6	Clinical remission	

**Table 27:** Disease activity states and relationship to DMARD escalation decisions

### Fulfilment of MSUS Indication Criteria by Assessment Group

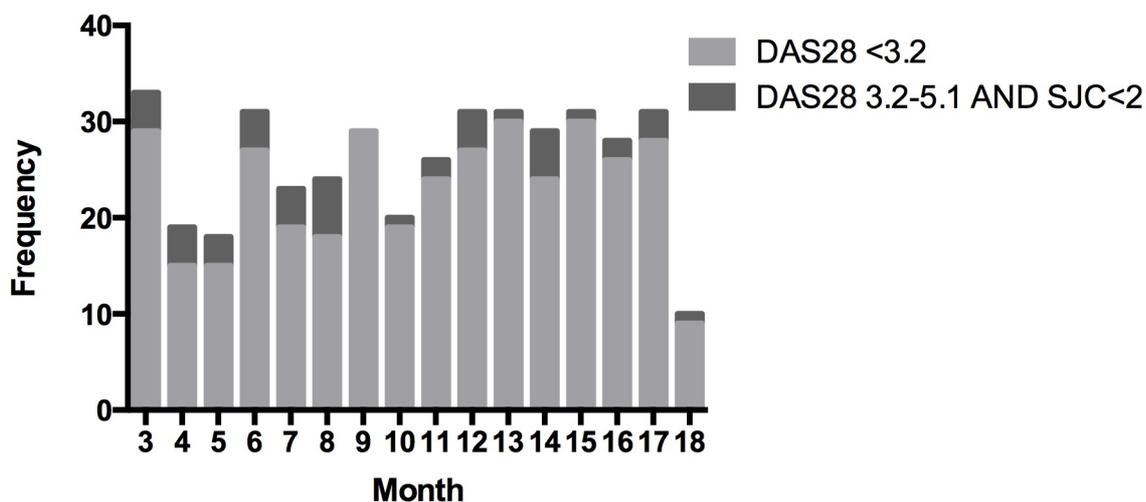
Between follow-up months 3 and 18, participants within the MSUS assessment group underwent 753 clinic consultations. Of these, 658 reviews fulfilled the pre-defined indications for MSUS disease activity assessment; on 580 occasions (88%) MSUS assessment was potentially indicated because DAS28 was less than 3.2 (LDAS – low disease activity) and on 78 occasions (12%) MSUS assessment was potentially indicated because DAS28 suggested moderate disease activity ( $3.2 \leq$  DAS28 < 5.1) with minimal clinical synovitis (28SJC $\leq$ 1). Similar rates of each of the different disease activity states were also identified in the whole cohort and in the DAS28 assessment group (Table 28). The most common clinical disease activity state was low disease activity and DAS28 remission was identified in over half of assessments in each group

Disease Activity State Descriptor	Total number of assessments Whole cohort	Number of assessments in MSUS group	Number of assessments in DAS28 Group
DAS28 $\geq$ 5.1	48 (3%)	12 (2%)	36 (5%)
3.2 $\leq$ DAS28 < 5.1 AND 28 SJC $\geq$ 2	156 (10%)	83 (11%)	73 (10%)
3.2 $\leq$ DAS28 < 5.1 AND 28SJC = 0-1*	177 (12%)	78 (10%)	99 (13%)
DAS28 < 3.2*	1139 (75%)	580 (77%)	449 (73%)
DAS28 < 2.6*	852 (56%)	432 (57%)	420 (55%)
Total number of assessments	1520	753	657

**Table 28:** Frequency (percentage) that DAS28 assessments fulfilled different disease activity states between 3 and 18 months.  
Data compiled from all visits for each assessment group  
\*MSUS indication

Altogether, up until February 2013, and allowing for necessary deferment of MSUS assessment after DMARD escalation, 414 MSUS assessments had been conducted on MSUS group participants. On 369 of these occasions (89%) MSUS assessment was indicated because of DAS28 LDAS and on 45 occasions (11%) MSUS assessment was indicated because of moderate disease activity and minimal clinical synovitis.

So far, relatively few sets of paired DAS28 and MSUS disease activity assessments are available for the 18 month time point since: 1. not all participants have completed the full follow-up period, 2. a number of participants had undergone DMARD escalation within the preceding 3 months and/or 3. because participants were about to leave the close monitoring processes allowed by study participation. Over the whole follow-up period DAS28 LDAS was consistently the most frequent indication for MSUS assessment (Graph 1). During months 3 to 12 of the follow-up period there was an apparent spike in the total number of MSUS assessments every 3 months that may represent the deferment of MSUS assessments for at least 3 months following DMARD escalation. Between months 12 to 17 the number of MSUS assessments being conducted each month became relatively static (between 28-31 MSUS assessments per month) with the majority being indicated by DAS28 LDAS.



Graph 1: Frequency of MSUS indication by month – MSUS assessment group only

## 5.2 Clinical Disease Activity Assessment Findings in MSUS Assessment Group

To determine whether the proposed MSUS indications were selecting appropriate participants for MSUS assessment the mean values of the DAS28 component variables were compared between the different DAS28 disease activity states. To avoid biasing the on-going management of continuing study participants a similar comparison was not performed for the DAS28 assessment group

### 5.2.1 Mean Values of DAS28 Component Variables

The available data for reviews of the MSUS assessment group between months 3 and 18 was segregated into groups based upon the previously defined disease activity states (Table 27) and the mean value of each of the DAS28 components for each group was calculated (Table 29). In summary, as the relative severity of the DAS28 disease activity assessment increased there was a corresponding increase in the individual components of the DAS28.

In the small number of occasions of high disease activity (DAS28>5.1) the individual mean DAS28 components clearly suggested active disease that was likely to be clinically evident. Patients exhibited the highest number of tender and swollen joints, reported the worst global health and pain scores and also exhibited the highest ESR. Conversely, instances of DAS28 LDAS and DAS28 remission were both associated with very low tender and swollen joint counts, suggesting that there was little clinical evidence of active synovitis. As a whole group, instances of moderate disease activity were associated with mean values that were somewhere between those of the low and disease activity groups. Dividing the moderate disease activity group using a threshold of 2 or more clinically swollen joints did identify differences between the mean values of the other DAS28 components. Compared to patients with 2 or more clinically swollen joints those patients with minimal clinical synovitis (SJC<2) tended to exhibit higher tender joint counts (mean: 3.55 vs 3.05), global health 10cm VAS (mean: 48.78 vs 35.33) and pain 10cm VAS (mean: 47.09

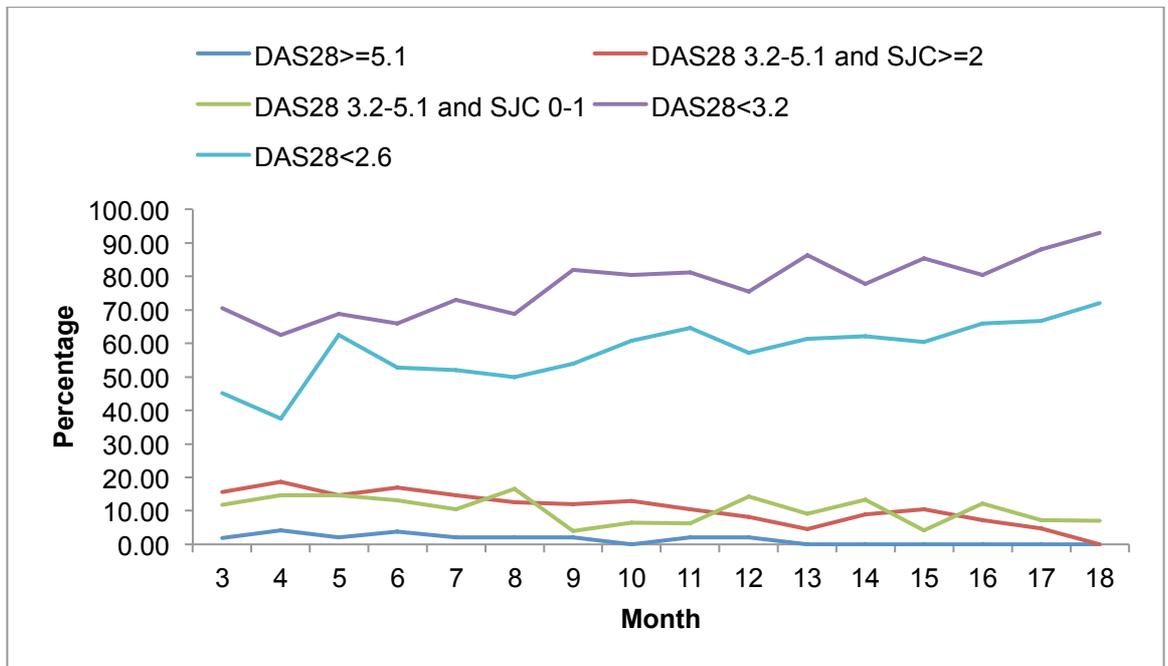
vs 40.45). However, even though there was no clinical evidence of synovitis the mean ESR was still moderately elevated providing some objective evidence of underlying active inflammatory disease

<b>DAS28 Disease Activity Assessment</b>	<b>DAS28</b>	<b>28TJC</b>	<b>28SJC</b>	<b>ESR</b>	<b>Global VAS</b>	<b>Pain VAS</b>
<b>DAS28&lt;2.6</b> N = 432	1.72 (0.49)	0.14 (0.44)	0.49 (0.92)	9.45 (6.24)	7.56 (12.36)	7.85 (13.37)
<b>DAS28&lt;3.2*</b> N = 580	2.02 (0.67)	0.36 (0.89)	0.69 (1.28)	12.50 (10.29)	11.27 (15.97)	11.84 (17.32)
<b>3.2≤DAS28&lt;5.1</b> N = 161	3.89 (0.52)	3.29 (2.88)	1.92 (1.88)	26.16 (20.22)	41.84 (25.42)	40.45 (25.30)
<b>3.2≤DAS28&lt;5.1 and SJC&lt;2*</b> N = 78	3.75 (0.48)	3.55 (2.95)	0.42 (0.50)	21.82 (14.46)	48.78 (23.07)	47.09 (22.22)
<b>3.2≤DAS28&lt;5.1 and SJC ≥2</b> N = 83	4.03 (0.53)	3.05 (2.82)	3.33 (1.60)	30.23 (23.80)	35.33 (25.92)	34.13 (26.55)
<b>DAS28 ≥5.1</b> N = 12	5.72 (0.57)	8.33 (5.03)	6.58 (4.08)	42.50 (25.53)	72.25 (17.37)	65.17 (20.24)

**Table 29:** Mean (SD) values of DAS28 components for MSUS assessment group participants between 3 and 18 months follow-up  
\* MSUS Indication

## **5.2.2 Frequency of Clinical Disease Activity States in MSUS Assessment Group by Month**

Over the course of the follow-up period there was a gradual shift in the proportion of participants classified as each of the different disease activity states (Graph 2). Overall, the proportion of participants with clinically active RA fell and was mirrored by a steady increase in the proportion with either DAS28 LDAS and/or remission. There were no instances of high disease activity from month 13 onwards. The proportion of participants with DAS28 LDAS rose from 71% at month 3 to 93% at month 18. Similarly, the proportion of participants with DAS28 remission rose from 45% at month 3 to 72% at month 18. The proportion of participants with moderate disease activity but minimal synovitis remained relatively static, suggesting that there may be a subset of RA patients in whom it is not possible to fully suppress DAS28.



**Graph 2:** Percentage occurrence of each disease activity state by month – MSUS assessment group

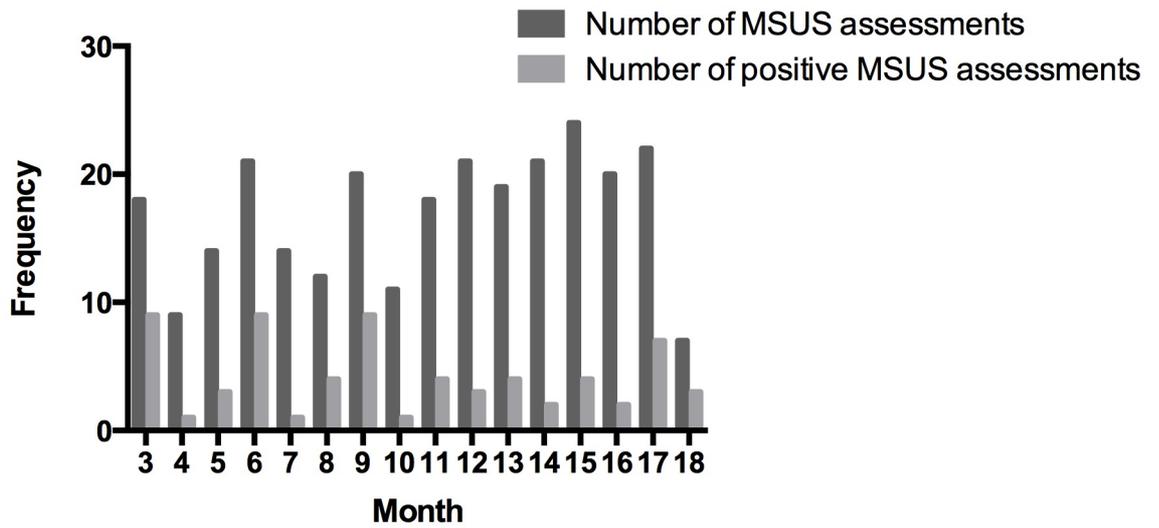
### 5.3 Potential workload implications of MSUS Assessment

To quantify the potential workload implications of incorporating MSUS assessment into routine clinical practice the incidence of each MSUS indication, the likelihood of each indication identifying active disease and the relationship between the frequency of MSUS assessments and the incidence of active disease for each patient was calculated.

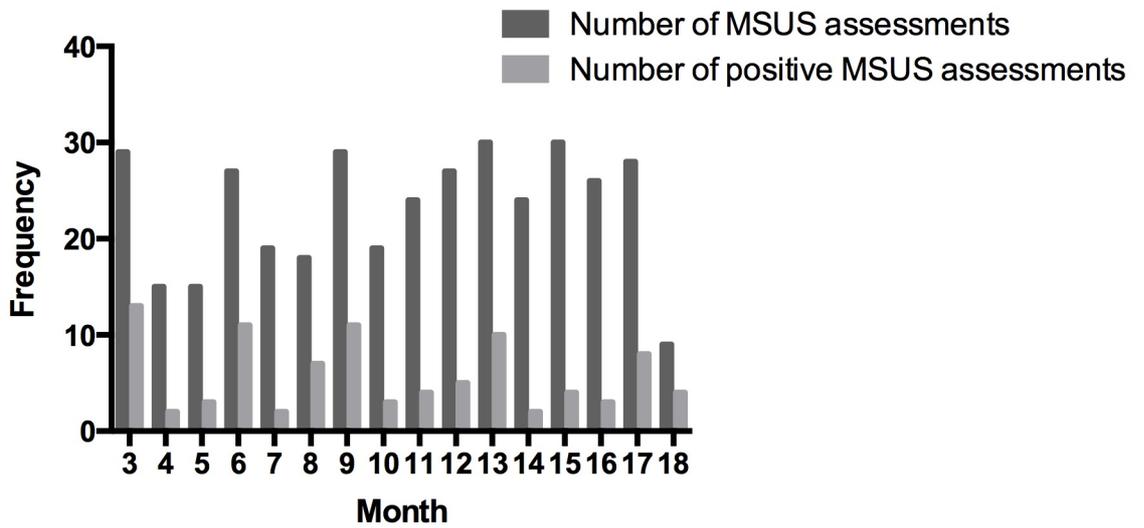
#### 5.3.1 Frequency of each MSUS indication and likelihood of identifying active disease

Regardless of indication, the majority of MSUS assessments failed to identify active synovitis. Three-hundred and sixty-nine MSUS assessments were conducted because of DAS28 LDAS; of these, 92 (25%) identified active disease. Similarly, 271 MSUS assessments coincided with DAS28 remission; of these, 66 (24%) identified active disease. Forty-five assessments were conducted because participant’s DAS28 suggested moderate disease activity but with minimal clinical synovitis; however, only 15 of these assessments (33%) actually confirmed the presence of underlying active synovitis.

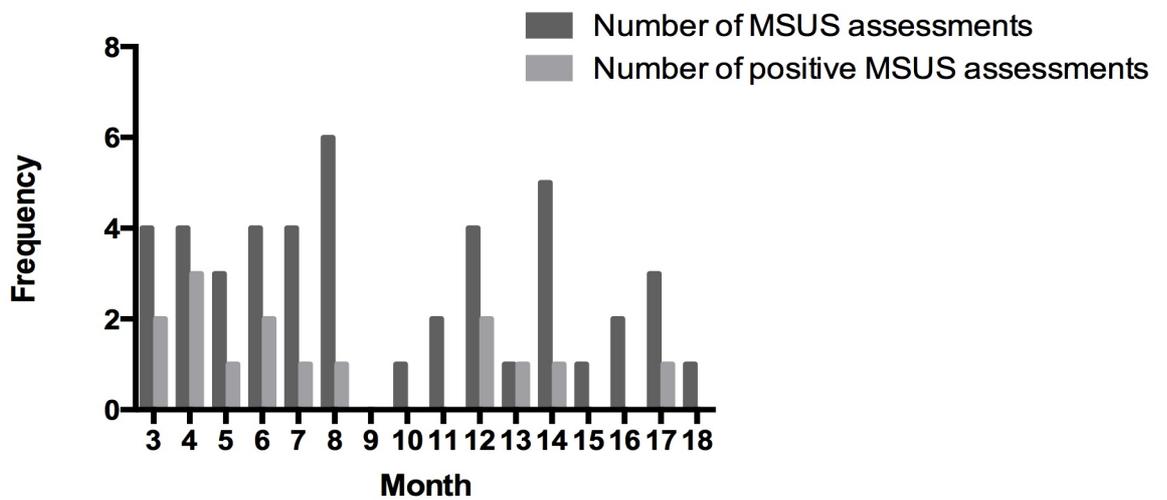
For each MSUS indication, there was a gradual fall in the total number of MSUS assessments that identified active disease over the follow-up period (Graphs 3-5). The frequency of positive MSUS assessments varied on a month-by-month basis (Graph 6). Further, there was also considerable variability evident when the proportion of positive MSUS assessments was segregated by indication (Graph 7). The highly variable nature of these findings may, in part, be explained by the relatively small number of MSUS assessments performed during some months.



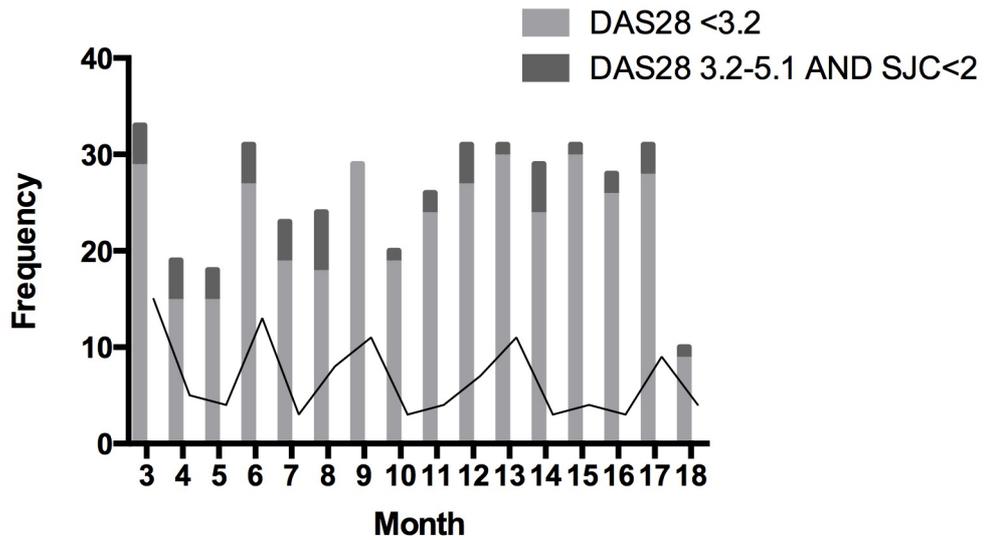
Graph 3 – DAS28<2.6 - Frequency of MSUS assessments and frequency of positive MSUS assessments



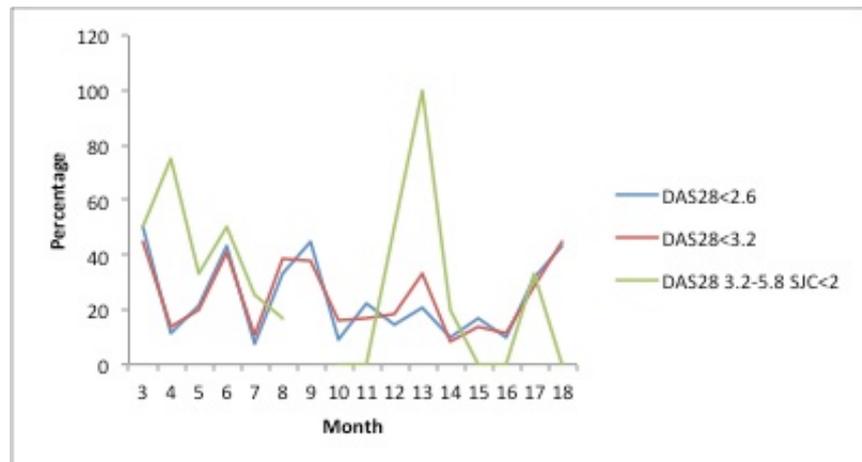
Graph 4 - DAS28 < 3.2 - Frequency of MSUS assessment and frequency of positive MSUS assessments



Graph 5 - 3.2≤DAS28<5.1 and SJC<2 – Frequency of MSUS assessment and frequency of positive MSUS assessments



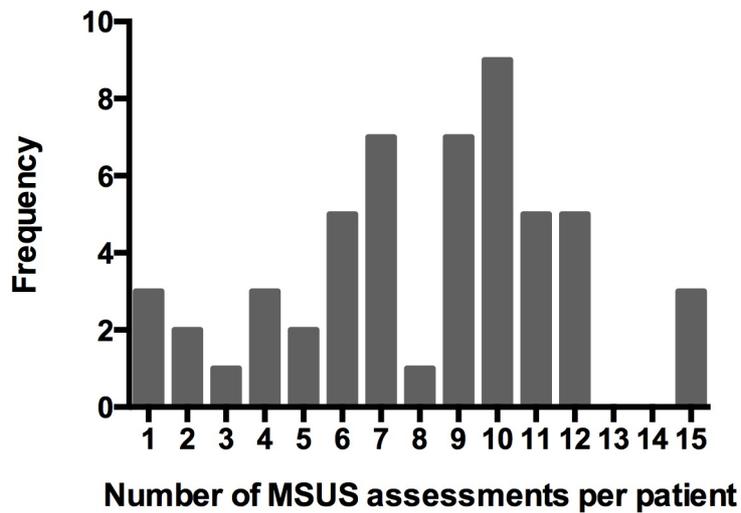
**Graph 6** – Relationship between frequency of MSUS assessment and frequency of positive MSUS assessments  
 line = frequency of positive MSUS assessment



**Graph 7** – Percentage of MSUS assessments that identified active disease by MSUS indication and month

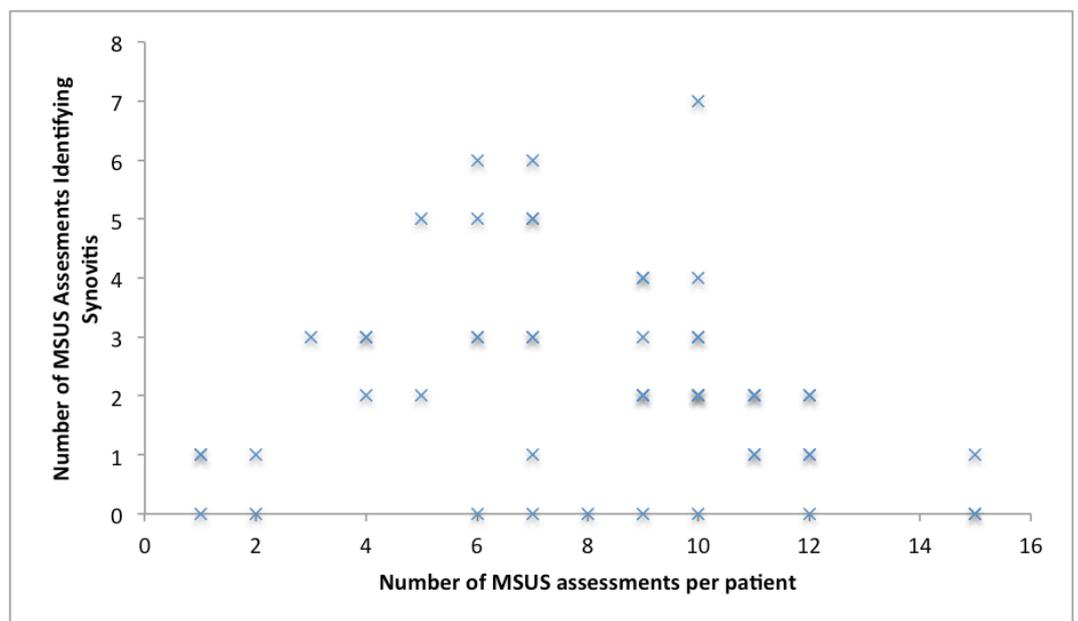
### 5.3.2 Relationship of frequency of MSUS assessment to incidence of active disease

All MSUS assessment group participants underwent at least 1 MSUS assessment over the follow-up period. The frequency of MSUS assessment for each patient demonstrated a positively skewed distribution (Graph 8). The median number of assessments performed per patient was 9 (IQR 6-10) whilst the mean was 8.2 (SD 3.5)



**Graph 8** – Distribution of number of MSUS assessments conducted per patient over follow-up period

The relationship between the number of the MSUS assessments performed per participant and the number which actually identified active disease was not linear. As Graph 9 demonstrates the proportion of MSUS assessments that actually identified active disease remained highly variable between individual participants. Notably, a small number of participants (n = 6) underwent comparatively infrequent MSUS assessments (between 1 and 3 assessments per participant) that identified a high rate of active disease (60% of assessments). It is presumed that this reflects the small subgroup of participants with highly active disease in whom MSUS assessment was only infrequently indicated since treatment escalation decisions could be based upon clinical findings alone. There was also a subset of participants (N = 13) who underwent comparatively high numbers of MSUS assessments (between 11 and 15 assessments per participant) that identified relatively low rates of active disease (percentage of positive assessments 2.2-14.5%) but still contributed significantly to the overall workload (160 MSUS assessments; 37% of total number).



**Graph 9** – Scattergraph demonstrating the relationship between the number of MSUS assessments performed per patient and the number of assessments that identified active disease

### 5.3.3 Indication and findings of first MSUS assessment

The majority of first MSUS assessments were conducted because participants exhibited DAS28 LDAS, rather than moderate disease activity and minimal clinical synovitis (Table 29). Twenty three (43%) of the first MSUS assessments identified active disease. However, despite the evident imbalance in numbers the proportion of MSUS assessments that identified active disease was similar for both MSUS indications (Table 30). Chi-squared analysis suggested that there was no statistically significant difference in the proportion of positive first MSUS assessments between either of the indication groups ( $p=0.8174$ )

Indication	Number (%)	Number (%) of positive MSUS assessments
DAS28<3.2	40 (76%)	17 (43%)
3.2≤DAS28<5.1 and SJC<2*	13 (24%)	6 (46%)
Total	53	23 (43%)

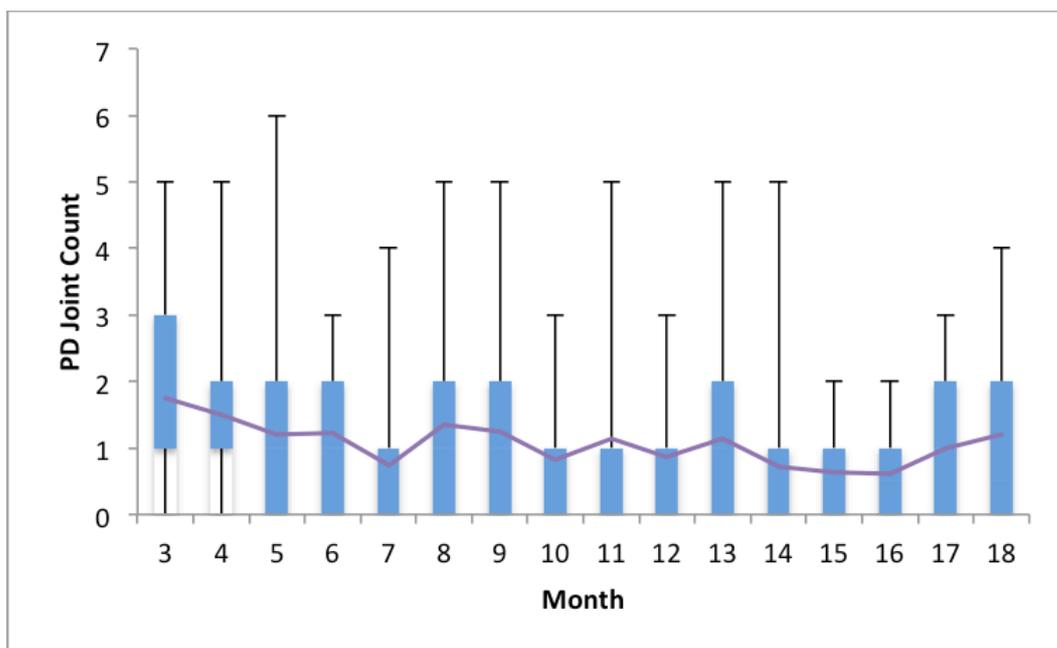
**Table 30** - Number of first MSUS assessments performed for each MSUS indication and number (percentage) that identified active disease

## 5.4 Change in MSUS PD signal findings over follow-up period

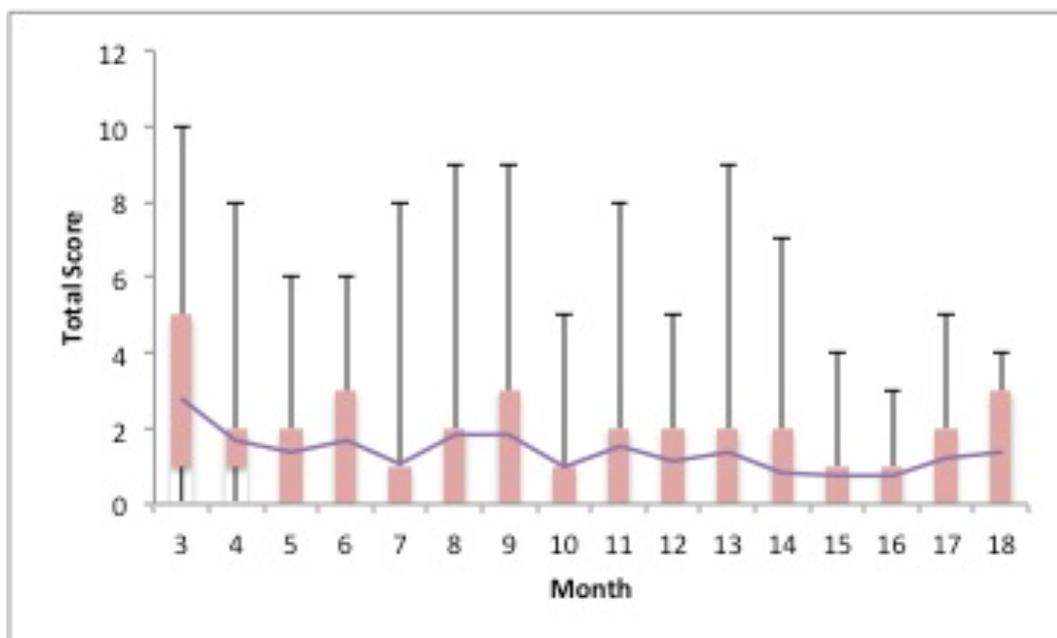
### 5.4.1 Overall change in MSUS findings

Despite the month-by-month variation in the rate of active disease identified by MSUS assessment there was a downward trend in the number and proportion of positive MSUS assessments over the whole of the follow-up period (Graph 6). A gradual fall in the total PD score and the number of joints demonstrating PD signal (the PD joint count) was observed, suggesting the total inflammatory burden was lessening and a positive treatment response. Mean PD score decreased from 2.70 to 1.34 (95% CI: 2.09, 0.63;  $p<0.001$ ) between the first and last MSUS assessment and PD joint count decreased from 1.78 to 1.12 (95% CI-1.10, -0.22;  $p=0.004$ ). Only 5 patients had undergone MSUS assessments at both month 3 and 18 time points. For these patients, a downward trend was observed in the change of mean total PD score (2.60 to 1.20) and mean PD joint count (1.80 to 1.00); however, due to the very small number of subjects, this did not reach statistical significance ( $p= 0.21$  and  $0.34$  respectively).

For each set of monthly MSUS assessments, the interquartile ranges of the PD joint counts fluctuated around a relatively narrow range of values (between 0-3) suggesting that the majority of MSUS assessments were identifying relatively small numbers of joints with evidence of PD signal (Graph 10). However, for each month, the range of all the PD joint counts observed was more variable than the relatively narrow band implied by the interquartile range. Even in the later follow-up months a subset (at least 25%) of participants continued to exhibit PD joint counts of 2 or higher that fulfilled DMARD escalation criteria. Therefore, continuing MSUS assessments up until this point remained worthwhile since it continued to identify evidence of active synovitis



**Graph 10** – Box and whisker plot of dispersal of PD joint counts recorded each month – all results.  
 Blue box – inter-quartile range; whiskers – range; purple line – mean



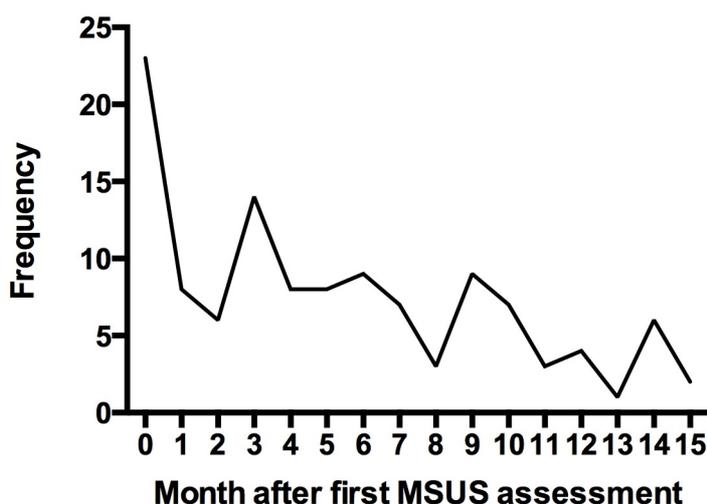
**Graph 11** – Box and whisker plot of dispersal of total PD score recorded each month  
 Pink box – inter-quartile range; whiskers – range; purple line – mean

Graphically representing the range of total PD scores demonstrates that there was a wide dispersal in the range of total scores and that some outlying patients continued to exhibit relatively high total PD scores, even at relatively late stages of follow-up (Graph 11). Once again, significant variability in the total range of scores was evident throughout most of the follow-up period. The inter-quartile ranges suggest that, on the majority of occasions, the majority of participants exhibited total PD scores within a narrow band of low PD burden (range 0-3 – i.e. at most 3 joints affected, each with grade 1 PD signal). However, between months 3 and 14 there were still outlying participants (up-wards whisker) exhibiting much higher total PD scores (range 3 – 10) suggesting a minority who experienced a much higher inflammatory burden. From month 15

onwards the absolute and range of values relating to the outlying participants falls and narrows markedly (range 1 – 5), perhaps suggesting that previously high total inflammatory burdens have started to recede in response to increasingly aggressive DMARD therapy regimens

#### 5.4.2 Time course assessment of changes in PD findings

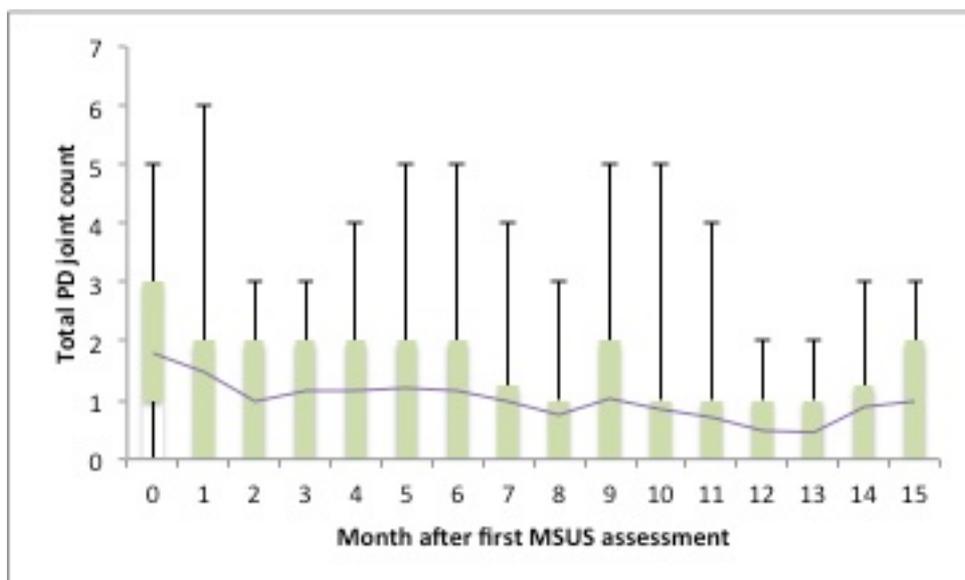
Since the pattern and sequence of MSUS assessments was highly variable between participants, describing the change in mean PD joint count on a month-by-month basis may not fully depict the impact of MSUS-steered DMARD therapy on MSUS findings since, for any given month, each participant will have been at a different place in the sequence of MSUS assessments. By taking the first MSUS assessment as a reference point, against which all subsequent MSUS assessments are compared, it is possible to describe how MSUS findings change once global disease activity starts to be assessed by a standardised MSUS measure. Following the first MSUS assessment, there was an overall decrease in the number of positive MSUS assessments performed during subsequent months (Graph 12: Month 0 - 23 positive assessments; Month 15 – 2 positive assessments,  $p=0.038$ ). However, it is worth restating that there were 53 data points available for the first MSUS assessment but only 5 for the fifteenth month after the first MSUS assessment.



**Graph 12** – Number of positive MSUS assessments observed for each month after first MSUS assessment

Following the first MSUS assessment there was a progressive fall in the total PD joint count between months 1 and 13 (mean PD count 1.8 vs 0.5 respectively,  $p<0.0001$ ), suggesting that the number of joints exhibiting any PD signal had also fallen (i.e. a positive treatment response) (Graph 13). However, after month 13 there was a slight increase in the mean PD joint observed to 0.9 (month 14) and then 1.0 (month 15). For the fourteenth month after first MSUS assessment there were 24 data points available, which is similar to preceding months 1 to 13 (range 22-32); therefore, it is likely that the apparent increase does represent a slight overall increase in the rate of active disease. However, for the 15<sup>th</sup> month there were only 5 data points available. Such discordance in data group sizes will have significantly increased the risk of single abnormal findings skewing the overall summary statistic and will have reduced the representative value of the month 15 data overall. Whilst the mean PD joint count was consistently less than 2, Graph 13

also demonstrates that at each time point there was a subgroup of participants (essentially the upper quartile) who exhibited significantly higher numbers of joints with PD signal



**Graph 13** - Total PD joint count score for each month after first MSUS assessment (month 0)  
Box = interquartile range, whiskers = maximum and minimum, line = mean

### **Determining the likelihood of future MSUS findings from preceding MSUS assessments**

Whilst the assessment and treatment protocol lead to participants undergoing a highly variable number of MSUS assessments, and usually for a mixture of indications, it also lead to many participants undergoing several consecutive negative MSUS assessments (Graph 9). By setting each participant's first MSUS assessment as a reference point, the relationship between different sequences of MSUS findings, and the likelihood of subsequently identifying active disease could be calculated. Several, theoretical, but clinically likely, scenarios were devised to determine the relationship between patterns of MSUS findings and the risk of identifying active disease during subsequent assessments. Active disease continued to be defined as at least 2 or more joints exhibiting any PD signal. Overall, a positive MSUS assessment was associated with an increased likelihood of subsequent MSUS assessments identifying active disease; whilst consecutive negative MSUS assessments were associated with a decreasing likelihood that subsequent MSUS assessments would identify active disease

*Probability of identifying active disease following a positive MSUS assessment – applies to two different clinical scenarios:*

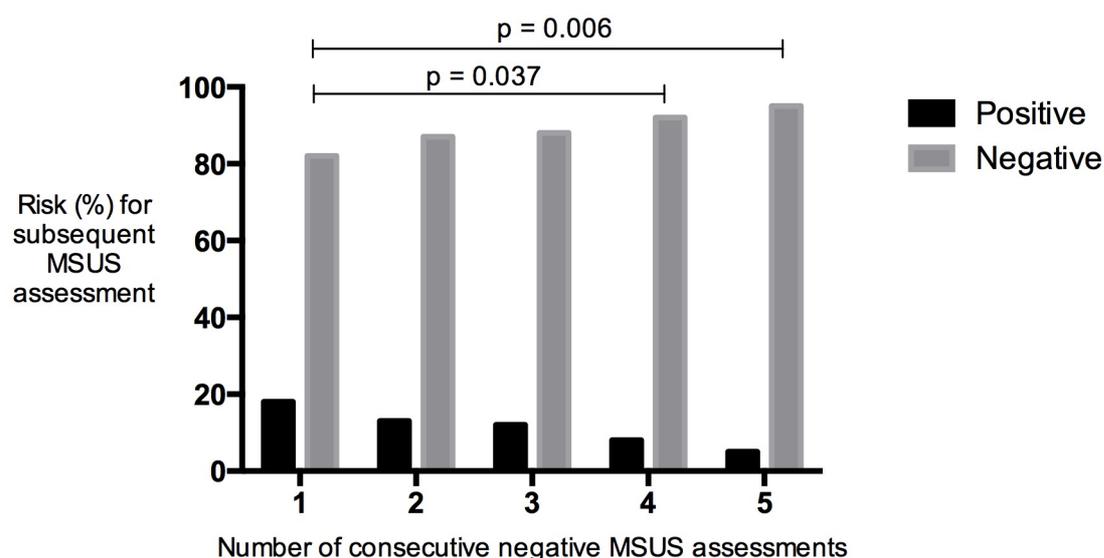
1. Three month follow-up MSUS assessment - if DMARD therapy had been escalated because MSUS assessment identified active disease it is likely that another MSUS assessment would not be performed for at least 3 months. Presuming that each positive MSUS assessment would have lead to DMARD escalation, the probability of any positive MSUS assessment being followed by another positive MSUS assessment three months later was calculated. Altogether, 55 positive MSUS assessments had 3 month follow-up

MSUS assessment data available. Of these, 30 follow-on assessments (55%) continued to identify active disease whilst 25 assessments (45%) did not

2. One month follow-up MSUS assessment – on a small number of occasions a positive MSUS assessment was not followed by immediate DMARD escalation; usually due to patient choice, intercurrent illness or the presence of a contra-indication (e.g. blood monitoring abnormality). If DMARD therapy was not escalated, a further MSUS assessment was conducted during the subsequent month's consultation to determine whether there was persistence of PD signal. Comparing the findings of both these assessments provided the opportunity to determine the likelihood of a single positive MSUS assessment being followed by another positive assessment one month later. There were 22 occasions when a positive MSUS assessment had follow on MSUS assessment data available for the next month. Eighteen follow-on assessments (82%) demonstrated persistence of active disease, whilst 4 follow-on assessments (18%) no longer identified active disease.

*Probability of identifying active disease following a negative MSUS assessment* – applies to two different clinical scenarios:

1. One month follow-up assessment – if a MSUS assessment did not identify active disease, the assessment protocol required that participant's global disease activity be reassessed during the next monthly review to determine whether they then fulfilled the criteria for DMARD escalation. There were 245 occasions when a single negative MSUS assessment had MSUS assessment data available for the next month. In 45 follow-up MSUS assessments (18%) active disease was identified and DMARD therapy would usually have been escalated. By contrast, 200 follow-up assessments (82%) showed no evidence of active disease and DMARD therapy would usually not have been changed
2. Risk of active disease after consecutive negative MSUS assessments – data from consecutive negative MSUS assessments was analysed to determine whether increasing numbers of consistently negative MSUS assessments were associated with an increased chance of future assessments also being negative (Graph 14). Overall, as the number of consecutive MSUS assessments increased the risk of a subsequent MSUS assessment identifying active disease gradually fell. Systematic between group comparisons (Fisher's test) demonstrated that, compared to one negative MSUS assessment, after 4 and 5 consecutive negative MSUS assessments there was a statistically significant reduction in the risk of subsequent MSUS assessments being positive ( $p = 0.037$  and  $0.0062$  respectively). All other between group comparisons did not reach statistical significance



**Graph 14** – Risk of subsequent MSUS assessment identifying active disease following consecutive negative assessments.

In summary, a positive MSUS assessment was associated with an increased likelihood of subsequent MSUS assessments identifying active disease; whilst a negative MSUS assessment was associated with an increased likelihood of subsequent MSUS assessments not identifying active disease. Compared to a single negative MSUS assessment the relative risk of a single MSUS assessment being followed by another positive assessment is 4.45 (95% CI 3.20 – 6.19;  $p < 0.0001$ ). Furthermore, increasing numbers of negative consecutive MSUS assessments were associated with a diminishing risk of subsequent MSUS assessments identifying active disease (Table 31)

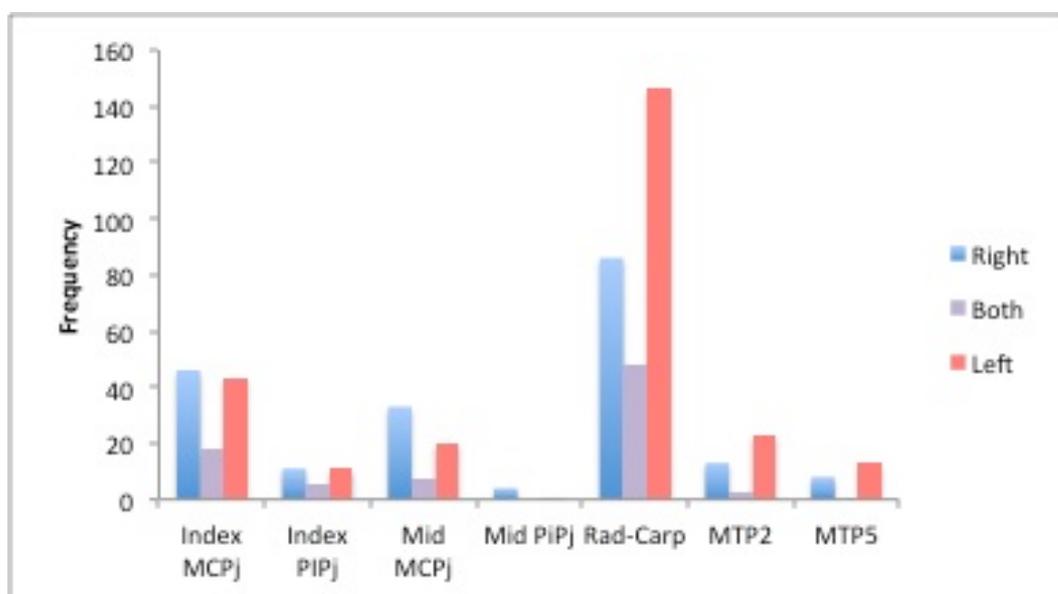
Preceding MSUS assessment findings	Relative risk of subsequent positive MSUS assessment	95% Confidence Intervals	p
Single positive assessment	4.45	3.20 – 6.19	< 0.0001
2 consecutive negative assessments	0.72	0.45 – 1.15	0.17
3 consecutive negative assessments	0.65	0.38 – 1.11	0.12
4 consecutive negative assessments	0.45	0.21 – 0.96	0.041
5 consecutive negative assessments	0.26	0.085 – 0.82	0.021

**Table 31** – Relative risk of subsequent MSUS assessments identifying active disease following consecutive negative MSUS assessments

## 5.5 MSUS PD signal findings by joint area

The available MSUS data were interrogated to determine how often each joint area displayed positive MSUS findings that contributed to the final treatment decision. The radiocarpal joint most frequently demonstrated PD signal. The index MCPj and middle MCPj were next most frequently involved. By comparison, the index PIPj, middle PIPj, second MTPj and fifth MTPj demonstrated PD signal very infrequently (Graph 15). The number of occasions when bilateral joint involvement was recorded was consistently less than the number of occasions when unilateral joint

involvement was identified. This asymmetrical pattern of joint involvement on MSUS is contrary to the traditional depiction of active RA as a symmetrical polyarthritis and may represent either very early disease or attenuation of the usual disease processes by DMARD therapy



**Graph 15 – Frequency and distribution of positive PD findings by joint area**  
Data presented for right (blue) and left (red) hand side involvement combines both unilateral and bilateral involvement of each joint area

### 5.5.1 Frequency of positive MSUS assessments using different joint sets

Graph 15 shows that, even using the limited 14 joint set, there were certain joint areas (notably radio-carpal and MCP joints) that were much more likely to exhibit positive PD signal. Equally, as the size of the joint set increases there will be a corresponding increase in the length of time needed for performing the MSUS assessment. Reducing the size of the joint set will reduce the overall assessment time; though may also reduce its sensitivity. Using the 14 joint set as reference, different combinations of joints were tested to determine what impact further reductions in the size of the joint set would have upon the overall accuracy of MSUS assessment (Table 32). To varying degrees, all of the reduced joint sets proved less sensitive than the 14 joint set. Joint sets that retained bilateral assessment of each joint area (sets 1 – 3) performed notably better than unilateral joint sets (sets 4 and 5). As the bilateral joint sets became more restricted the accuracy of the disease activity assessment also fell. For example, by omitting the 2<sup>nd</sup> and 5<sup>th</sup> MTP joints the overall sensitivity of assessment fell to 90% (set 2). By contrast, retaining the 2<sup>nd</sup> and 5<sup>th</sup> MTP joints but omitting the index and middle PIP joints (set 3) had the smallest impact on overall sensitivity (97%). Omitting both the index and middle PIP joints and the 2<sup>nd</sup> and 5<sup>th</sup> MTP joints caused the sensitivity to fall to 86.7%. Thus, the 2<sup>nd</sup> and 5<sup>th</sup> MTP joints, which individually demonstrated slightly higher rates of involvement than the index and middle PIP joints, also had a greater impact on the likelihood of diagnosing active disease

Joint Set Name		Active Disease (PD Signal 2+ joints)	Inactive Disease (PD Signal 0-1 joints)	Sensitivity
<b>TaSER Set</b>	14 joints – bilateral RCj, MCP2, MCP3, PIP2, PIP3, MTP2, MTP5	118 (28%)	311 (72%)	100%
<b>1</b>	6 joints – bilateral RCj, MCP2 + MCP3	100 (23%)	329 (77%)	86.7%
<b>2</b>	10 joints – bilateral RCj, MCP2, MCP3, PIP2 + PIP3	105 (24%)	324 (76%)	90.0%
<b>3</b>	10 joints – bilateral RCj, MCP2, MCP3, MTP2 + MTP5	114 (27%)	315 (73%)	97%
<b>4</b>	All right hand sided joints	33 (8%)	396 (92%)	58%
<b>5</b>	All left hand sided joints	43 (10%)	386 (90%)	61%

**Table 32** – Frequency (percentage) of positive MSUS assessments using different joint sets

## 5.6 Agreement between DAS28 and MSUS disease activity assessments

### 5.6.1 Definitions (reprise)

Definitions of agreement (and disagreement) between DAS28 and MSUS disease activity assessments was based on whether the outputs of each assessment method would have resulted in the same decision to intensify (or not) DMARD therapy. Agreement was present if both methods would have produced the same treatment decision (i.e. they agreed on the disease activity state) and disagreement was present if both methods would have produced opposing treatment decisions (i.e. disease activity state assessments were discordant). Table 28 (reprinted below), summarises the interaction between DAS28 and MSUS disease activity assessments in relation to DMARD escalation decisions

Disease Activity State Descriptor	Disease State	DMARD escalation decision
DAS28 $\geq$ 5.1	High disease activity	Escalate on clinical grounds
$3.2 \leq$ DAS28 $<$ 5.1 AND 28SJC $\geq$ 2	Moderate disease activity and clinical synovitis	
$3.2 \leq$ DAS28 $<$ 5.1 AND 28SJC = 0-1	Moderate disease activity but minimal clinical synovitis	MSUS assessment required to inform DMARD decision
DAS28 $<$ 3.2	Low disease activity	
DAS28 $<$ 2.6	Clinical remission	

**Table 28:** Disease activity states and relationship to DMARD escalation decisions

### **5.6.2 MSUS disease activity assessment findings during DAS28 LDAS**

Three hundred and sixty-nine MSUS assessments were performed during DAS28 LDAS, of these 277 (75%) showed no evidence of active disease and therefore supported the impression given by DAS28 that little, or no, active disease remained. Ninety two (25%) MSUS assessments disagreed with the DAS28 by identifying PD signal in 2 or more joints (Table 33).

If the MSUS PD joint count DMARD escalation threshold was reduced to include assessments that identified PD signal in any joint, 219 MSUS assessments (59%) would have suggested active disease (Table 34). Alternatively, if the threshold had been raised to only include occasions when 3 or more joints exhibited PD signal, 33 (9%) of the MSUS assessments would have been classified as active disease (Table 35).

### **5.6.3 MSUS disease activity assessment findings during DAS28 remission**

Two hundred and seventy-one MSUS assessments coincided with DAS28 remission. Two hundred and five (76%) assessments did not identify ultrasonographic evidence of active disease and therefore did not influence DMARD therapy. Sixty-six (24%) assessment did identify ultrasonographic evidence of active disease and supported further DMARD escalation even though participants fulfilled DAS28 remission criteria (Table 33).

Using an MSUS PD joint count threshold of one, 162 assessments (60%) would have been classified as showing active disease (Table 34). However, if the PD joint count threshold had been raised to 3, only 25 MSUS assessments (9%) would have identified active disease (Table 35) and the degree of agreement (and the number of MSUS assessments classified as inactive disease) would have appeared far higher

### **5.6.4 MSUS disease activity assessment findings during moderate disease activity and minimal clinical synovitis**

Forty-five MSUS assessments were performed when  $3.2 \leq \text{DAS28} < 5.1$  (moderate disease activity) and  $\text{SJC} < 2$  (minimal clinical synovitis) to determine whether the elevated DAS28 was associated with underlying active synovitis. Thirty (67%) MSUS assessments did not identify evidence of active disease and prevented further DMARD escalation on that occasion. Fifteen (33%) MSUS assessments confirmed the presence of active disease and supported the DAS28-based decision to escalate DMARD therapy (Table 33).

Lowering the MSUS PD joint count threshold to 1 would have classified 25 MSUS assessments (56%) as active disease (Table 34). Conversely, a MSUS PD joint count threshold of 3 or more would have only classified 6 instances (13%) as active disease (Table 35)

		MSUS Assessment	
		PD signal in 0-1 joints	PD in 2 or more joints
DAS28 Assessment (N = 414)	DAS28 <2.6 (N= 271)	205 (76%)	66 (24%)
	DAS28 <3.2 (N= 369)	277 (75%)	92 (25%)
	3.2 ≤ DAS28 <5.1 and SJC <2 (N= 45)	30 (67%)	15 (33%)

**Table 33** – Agreement (green boxes) and disagreement (red boxes) between DAS28 and MSUS assessments of global disease activity.  
MSUS threshold = PD signal in 2 or more joints  
PD – Power Doppler, SJC – swollen joint count

		MSUS Assessment	
		PD signal in 0 joints	PD in 1 or more joints
DAS28 Assessment (N = 414)	DAS28 <2.6 (N= 271)	109 (40%)	162 (60%)
	DAS28 <3.2 (N= 369)	150 (41%)	219 (59%)
	3.2 ≤ DAS28 <5.1 and SJC <2 (N= 45)	20 (44%)	25 (56%)

**Table 34** – Agreement (green boxes) and disagreement (red boxes) between DAS28 and MSUS assessments of global disease activity.  
MSUS threshold = PD signal in 1 or more joints

		MSUS Assessment	
		PD signal in 0-2 joints	PD in 3 or more joints
DAS28 Assessment (N = 414)	DAS28 <2.6 (N= 271)	246 (91%)	25 (9%)
	DAS28 <3.2 (N= 369)	336 (91%)	33 (9%)
	3.2 ≤ DAS28 <5.1 and SJC <2 (N= 45)	39 (87%)	6 (13%)

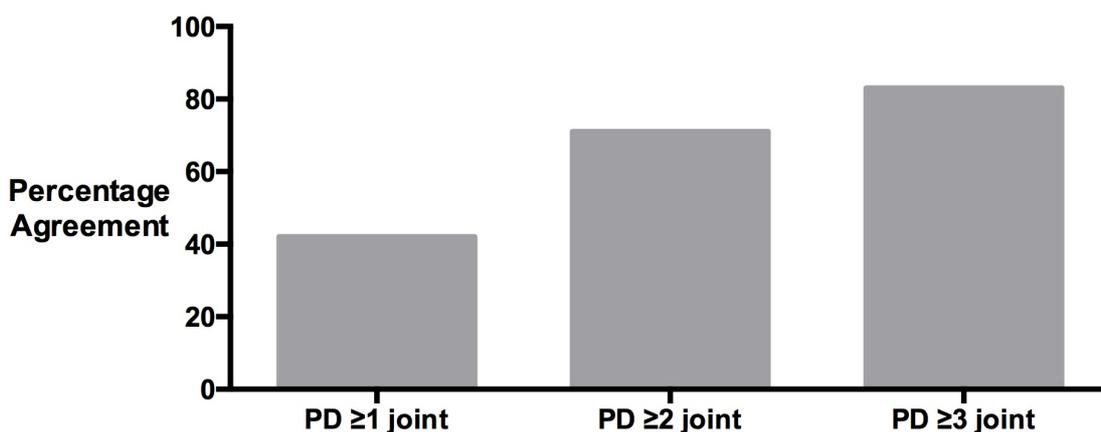
**Table 35** – Agreement (green boxes) and disagreement (red boxes) between DAS28 and MSUS assessments of global disease activity.  
MSUS threshold = PD signal in 3 or more joints

### 5.6.5 Overall agreement between DAS28 and MSUS disease activity assessments

The degree of overall agreement between DAS28 and MSUS disease activity assessment was dependent upon the MSUS PD joint count threshold used to define active disease (Graph 16). Using the proposed joint count threshold of 2 joints, there was agreement between DAS28 and MSUS disease activity assessments on 292 (71%) occasions. This also infers that MSUS disease activity findings had the potential to alter DMARD escalation decisions on 122 (29%) occasions. On 92 (22% of all MSUS assessments) of these occasions, MSUS assessment lead to DMARD therapy intensification that had not been suggested by DAS28 and on 30 occasions (7%) MSUS

assessment prevented DMARD escalation even though DAS28 registered moderate disease activity

A MSUS PD joint count threshold of 1 or more lead to more participants being classified as having active disease but a lower overall rate of agreement (42%) with DAS28 (Graph 16). Conversely, a MSUS PD joint count threshold of 3 or more lead to far fewer participants being classified as having active disease by MSUS assessment but an increased overall rate of agreement (83%).

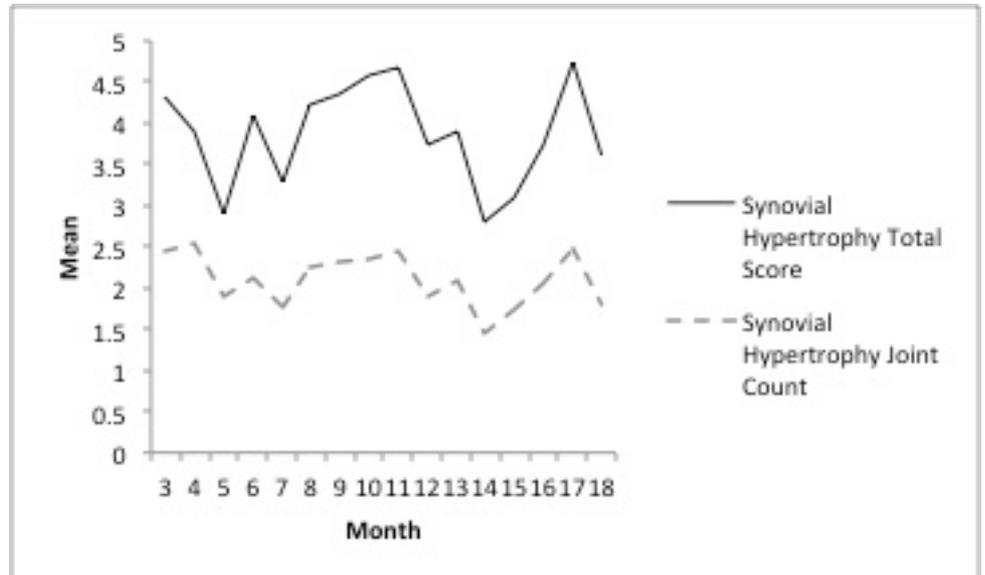


**Graph 16** - Percentage agreement between DAS28 and MSUS disease activity assessments based upon different MSUS PD joint count thresholds

## **5.7 Change in MSUS synovial hypertrophy findings over follow-up period**

The proposed definition of active disease focuses on the presence of intra-articular PD signal as the primary MSUS marker of active synovitis. Assessing the presence of PD signal also allows direct visualisation of gray scale synovial hypertrophy, and, even though these findings did not influence DMARD therapy decisions directly, the available data does make it possible to describe how synovial hypertrophy appearances changed over the follow-up period.

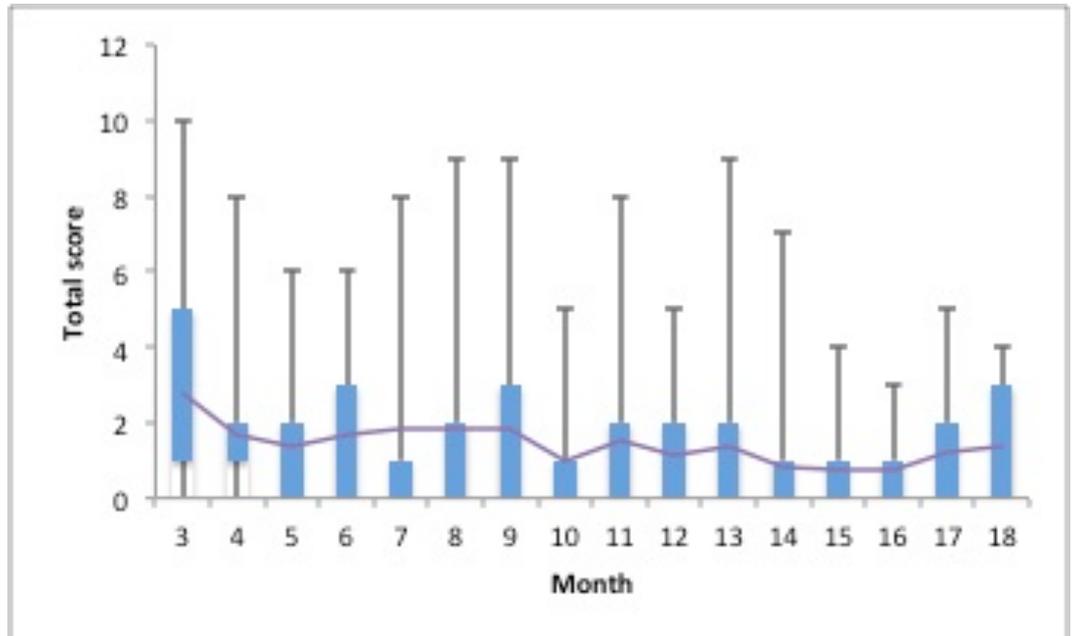
Whilst an overall downward trend was observed in both the total PD score and joint count over the follow-up period, the pattern and degree of change in the mean total synovial hypertrophy score and joint count was much less marked. Considerable month-by-month variability was observed in the mean total synovial hypertrophy score and joint count and in fact there was actually very little overall change (Graph 17). The greatest degree of fluctuation in both total synovial hypertrophy score and joint count is evident between follow-up months 11 and 18. Interestingly, excepting month 18, this is also the period when the number of MSUS assessments performed each month is most consistent (28 – 32 assessments per month). Therefore, whilst presumably there will have been some low level variation in which specific participants underwent MSUS assessment each month, it is unlikely that the observed synovial hypertrophy findings will have been excessively skewed by small numbers of aberrant outliers.



**Graph 17** – Mean total synovial hypertrophy score (black line) and joint count (dotted line) by month

Overall, a small numerical fall was observed in the mean total synovial hypertrophy scores. However, the degree of change did not meet statistical significance for comparisons comprising the mean first (4.58) and last (4.34) synovial hypertrophy score (mean difference -0.24; CI -1.23 – 0.75; p 0.63) or the month 3 (3.20) and month 18 (2.00) values (mean difference -1.20; CI -5.96, 3.56; p 0.52). Thus, whilst step-up DMARD therapy was associated with an apparent overall improvement in total PD score over the follow-up period it was not associated with a similar improvement in the total synovial hypertrophy score. Graph 18 demonstrates that the total synovial hypertrophy score recorded each month was also variable for an outlying subset of participants. In the majority of participants (blue box) the total synovial hypertrophy score fell within a relatively narrow band (0-3). However, up until month 13, MSUS assessments continued to record much higher total synovial hypertrophy scores in a subset (approximately 25%) of participants. Interestingly, from month 14 onwards, there is an apparent, and consistent, decrease in the maximum score recorded and a narrowing in the total range of scores. This may reflect prolonged DMARD therapy being eventually associated with some reduction in the total volume of synovial hypertrophy present in outlying participants; however, it may not yet be sufficient to influence the overall group mean. The follow-up period would need to be extended to determine whether the apparent reduction in score continued beyond month 18.

A small, non-significant fall was also observed in the mean synovial hypertrophy joint counts over the follow-up period. Between the first and the last recording mean synovial hypertrophy joint count fell from 2.48 to 2.30 (mean change -0.18; CI -0.74 – 0.38; p 0.52). Similarly, between the month 3 and month 18 time points, mean synovial hypertrophy joint count fell from 1.60 to 1.00 (mean change -0.60, CI -3.02 – 1.82; p 0.53).



**Graph 18** – Box and whisker plot of range of values for total synovial hypertrophy score for each month  
 Blue box – interquartile range; whiskers – range; purple line – mean

## 5.8 Discussion

The available results do suggest that a standardised MSUS assessment of global disease activity will alter DMARD treatment escalation decisions in a significant subset of patients. In some instances of DAS28 LDAS and clinical remission (approximately 25%), MSUS assessment will support further DMARD escalation in patients who continue to exhibit evidence of subclinical synovitis. Conversely, in a significant proportion of patients with moderate disease activity, but minimal clinical synovitis (approximately 67%), MSUS assessment will support non-escalation of DMARD therapy in those who exhibit no ultrasonographic evidence of active disease. When considering how feasible, or appropriate, it would be to incorporate MSUS assessment into the routine assessment of RA patients a number of important factors should be considered:

1. How to select appropriate patients for MSUS assessment
2. What constitutes an adequate MSUS assessment of global disease activity and how best to define 'active disease' using MSUS findings
3. The additional workload that would be created by regularly performing the MSUS assessment
4. How frequently the MSUS assessment lead to altered treatment decisions
5. Whether the treatment decisions that were influenced by MSUS assessment lead to meaningful improvements in clinical, functional and/or radiological outcomes

### 5.8.1 Selection of patients for MSUS assessment

For logistical reasons it would be impractical, and illogical, to perform MSUS disease activity assessment upon all RA patients during every single review. In order to balance the time required to perform MSUS disease activity assessment against the needs of a busy clinical service MSUS assessment should be targeted at those patients who are most likely to benefit from its findings. Therefore, there will be some situations in which MSUS assessment is superfluous since

decisions to alter DMARD therapy can be supported by clinical findings alone. Conversely, there will also be some clinical situations where clinical examination is insufficiently sensitive and/or specific to provide an accurate reflection of the actual disease burden and could lead to inappropriate therapeutic decision making. Thus, for this study, the investigators devised clinical scenarios where the additional disease activity information provided by MSUS assessment was likely to alter the DMARD escalation decision. Considering the disease activity variables of each of these situations individually will determine whether appropriate indications for MSUS were proposed:

#### *DAS28 LDAS - including DAS28 remission*

Performing MSUS disease activity assessment during instances of DAS28 LDAS (with or without DAS28 remission) does seem reasonable because, at a group level, a notable number of assessments (approximately 25%) did identify evidence of active synovitis and supported further DMARD escalation. Equally, the summary values of the DAS28 components for instances of LDAS equated to a previously published definition of RA minimal disease activity (295,329). During these instances, the individual DAS28 components suggest that synovitis would not have been clinically evident. Therefore, MSUS assessment could provide an additional means of differentiating between DAS28 LDAS patients with no active synovitis (i.e. a truer definition of remission) and those who still have subclinical synovitis and might require further DMARD escalation. DAS28 LDAS was by far the most common indication (89%) for MSUS assessment; therefore a MSUS indication based upon DAS28 LDAS could lead to repeated negative assessments of patients with stable disease control. Following 4 or 5 consecutive negative MSUS assessments there was a low risk that further MSUS assessments would identify active synovitis therefore, it seems reasonable to stop performing MSUS after 4 consecutive negative assessments

#### *DAS28 Moderate disease activity ( $3.2 \leq \text{DAS28} < 5.1$ )*

Until recently, any instance of DAS28 moderate disease activity was considered to reflect ongoing active synovitis. However, it is increasingly recognised that, in some instances, other coexistent pathologies may also contribute to elevations in DAS28 scores that can be misinterpreted as being RA related (154,155,204). If there is evidence of clinically swollen joints the clinician can be confident that at least some of the elevation in DAS28 is related to active RA because of the presence of clinical synovitis. If there is minimal clinical synovitis, the clinician may consider that alternative, non-inflammatory processes are contributing to the DAS28 score, or may need to seek additional evidence to support further DMARD escalation.

The findings of this study support the understanding that the presence of clinically swollen joints is an important (but not absolute) indicator of active disease during DAS28 moderate disease activity. Accepting that only a limited joint set was examined by MSUS, this study suggests that the majority (67%) of instances of DAS28 moderate disease activity but minimal clinical synovitis were in fact not associated with evidence of active subclinical synovitis. Therefore, in these instances relentlessly escalating immunomodulatory therapy purely based on the value of the DAS28 may not be appropriate since alternative processes may be contributing to the patient's

on going symptoms. However, participants with DAS28 moderate disease activity but minimal clinical synovitis were still clearly symptomatic since they reported notable elevations in global health and pain 10cm VAS. They also had mild elevations of ESR, suggesting underlying inflammatory disease. Further, there were still a notable number of instances (33%) when MSUS assessment did identify active disease and supported the decision to escalate DMARD therapy. So the absence of clinically swollen joints doesn't totally exclude the presence of synovitis. Thus, during instances of moderate disease activity with minimal clinical synovitis, MSUS disease activity assessment does allow treatment decision to be tailored to the specific needs of the participants. If active synovitis is confirmed, symptomatic patients will still be considered for further DMARD escalation. By contrast, symptomatic patients who have no synovitis will avoid unnecessary (and potentially toxic) additional DMARD therapy but may benefit from being considered for alternative treatment approaches

#### *High disease activity (DAS28>5.1)*

Compared to other disease activity states there were comparatively fewer instances of high disease activity and none that underwent MSUS assessment. Further, the mean values of the individual DAS28 component variables suggested a clinical phenotype that is easily recognisable as active RA. Therefore it is unlikely that MSUS assessment would have added sufficient additional disease activity information to lead to a decision not to escalate DMARD therapy. Previous studies have reported that patients with fibromyalgia (204) and fibromyalgic-RA (155) do exhibit DAS28 scores that are significantly higher than other RA patients and in some instances reach DAS28 high disease activity. During this study, concerns relating to non-specific elevations of the DAS28 were focussed on the moderate disease activity band and all instances of high disease activity were considered RA related. Clearly, some of the instances of high disease activity may have been related to non-RA processes; however, since MSUS assessment was not performed during these instances the exact extent cannot be quantified. A significant overlap with fibromyalgia seems unlikely, since: 1. all instances of high disease activity were associated with at least 2 clinically swollen joints and 2. the tender and swollen joint counts were relatively well matched which is contrary to the description of fibromyalgic-RA proposed by Pollard et al (155)

Overall, the proposed indications do seem to select populations of patients that are appropriate for MSUS assessment of global disease activity. In particular, focussing on patients with DAS28 LDAS or DAS28 moderate disease activity but minimal clinical synovitis identified clinical situations where the additional MSUS findings contributed meaningfully to tailored treatment decisions during a significant subset of consultations. Furthermore, restricting MSUS assessment to occasions when there is clinical doubt about the disease activity state will avoid needing to perform potentially time consuming MSUS assessments during instances where there is clear clinical evidence of active synovitis. The results also suggest that it will be possible to stop performing MSUS assessment after 4 or 5 consecutive negative assessments since the risk of a subsequent assessment being positive is very low.

The proposed indications for MSUS assessment are based heavily on the consensus DAS28 definitions of disease activity. These definitions are relatively arbitrary and don't always provide a

full representation of the true disease state. Using the DAS28 patients can exhibit high numbers of swollen joints (i.e polyarticular synovitis) but still be classified as DAS28 LDAS or remission. As an extreme illustration: if 28SJC = 28, 28TJC = 0, ESR = 5 and global VAS = 0, DAS28 = 2.61. Indeed, as has previously been shown, a significant proportion of patients who meet DAS28 definition of remission can still demonstrate evidence of synovitis on imaging studies (121) and patients with DAS28 remission who still exhibit clinically swollen joints also exhibit progressive radiographic damage (192). Furthermore, it is possible that some RA patients who achieve high disease activity might have additional non-RA related processes adding to their overall DAS28 score. Nevertheless, the DAS28 score, and its shortcomings, is easy to apply, validated and well understood by most rheumatologists (272). If the presence of clinically swollen joints, and by extension clinical synovitis, is predictive of future radiographic progression an alternative approach that disregards the total DAS28 value and determines whether patients require MSUS assessment based upon their swollen joint count might be more appropriate. This alternative approach is categorical rather than linear and favours the identification of active synovitis by either clinical and/or MSUS assessment rather than the attainment of a predefined DAS28 value. MSUS assessment would then be indicated if the clinician felt there was minimal clinical synovitis present, or sufficient doubt relating to the clinical findings. For example, patients with DAS28 remission but 2 clinically swollen joints would be classified as active disease, as would patients with moderate-high disease activity if they exhibited at least 2 clinically swollen joints; whereas, patients with one or no clinically swollen joints would require MSUS assessment regardless of the DAS28 score.

### **5.8.2 Proposed MSUS joint set and definitions of active RA**

Whilst there is a mounting momentum of evidence arguing for the inclusion of some form of musculoskeletal ultrasound examination into the routine assessment of global disease activity in RA (129,330), there is continuing debate about which joints should be examined to assess global disease activity and what is the minimal level of joint involvement required to define active synovitis on ultrasonographic grounds. For the purposes of this research, a pragmatic joint set and MSUS definition of active disease were chosen:

#### *Proposed MSUS limited joint set*

So far, limited MSUS joint sets have been proposed comprising 7 joints (290-292), 12 joints (226) and 20 joints (293). Further, limited joint sets have been shown to have similar metric properties to more extensive sets (226,293,331), sensitivity to changes in disease activity (226,292,332) and simplified PD joint sets have demonstrated similar sensitivity to sets that also include grey scale synovial hypertrophy (226). Since there is no universally accepted joint set, a novel joint set was devised that pragmatically combined the features of previously proposed sets to be in keeping with the defining clinical characteristics of RA.

Unlike the studies by Dougados et al(293) and Naredo et al(226), the proposed MSUS joint set was devised by deduction rather than being derived from the findings of an extensive joint set. Thus, it is not possible to report how well the reduced set would perform in relation to a more extensive joint set. The concern being that by omitting particular joints that frequently exhibit

evidence of ultrasonographic synovitis the sensitivity of the proposed joint set is reduced. Equally, since the proposed MSUS joint set incorporates a novel combination of joints it is not possible to be certain that it will have similar metrological properties to any of the previously published joint sets. In order to standardise the examination technique, the joint set was non-adaptive; that is, symptomatic areas outwith the joint set were not examined. This is unlikely to represent usual clinical practice where clinicians may consider scanning all symptomatic areas a more appropriate assessment of global disease activity. Future studies in a similar area may increase the sensitivity of the MSUS assessment, and therefore its effectiveness, by considering either a much more extensive joint set (which requires longer scanning times and greater MSUS expertise) or a limited joint set and any other symptomatic region. This latter approach is particularly interesting since a major advantage of MSUS over MRI is its ability to examine most joint regions using a single piece of equipment.

The high incidence of PD signal identified at the wrist is striking when compared to other joints and will have contributed to a significant number of DMARD escalation decisions. Compared to other joints within the set, the wrist is significantly more complex anatomically, since it comprises articulations between the radius, ulnar and both carpal rows. Hence there is a greater risk that variability in MSUS technique will influence interpretation of synovial findings. To minimise inconsistencies related to MSUS technique, assessment at the wrist was limited to the radiocarpal joint, which is relatively easy to identify and examine in a standardise manner. Clearly, there may have been some instances when the adjacent inter-carpal joint displayed PD signal that was not evident at the radio-carpal joint and would have been recorded (perhaps incorrectly) as a negative assessment. In future studies, the presence of synovitis at either the radio-carpal, and/or the inter-carpal joint, may be a more appropriate overall assessment for wrist synovitis.

As the frequency of PD signal at the wrist was much higher than in other joint areas it is also important to consider whether findings at the wrist represented true synovial pathology. Until recently the assumption was that any intra-articular PD signal represented underlying synovial vascularity and, by extension, synovitis (213). However, it is possible that some PD findings are in fact related to the increasing sensitivity of modern ultrasound machines. Recent observational studies have reported a particularly high incidence (between 56-86%) of Doppler signal in the wrist joints of healthy volunteers but a much lower incidence in the MCP (7-11%) and PIP (1-2%) joints (224,295). This higher rate of Doppler signals may reflect the wrist having a much larger synovial surface than other peripheral joints and may be more likely to contain larger nutrient vessels that are detectable during Doppler examination. Importantly, failure to differentiate Doppler signal related to synovitis from that which is physiological (or artefactual) could contribute to inaccurate classification of the overall disease activity state and an inappropriate decision to escalate therapy. The resistive index (the degree of flow in tissue during diastole) can be used to differentiate physiological flow in synovial tissue from that related to inflammation; non-inflamed tissues exhibit a high resistance to flow, whilst inflamed tissues exhibit a comparatively lower resistance (295,333,334). Unfortunately, the resistive index was not measured, therefore, it is not

possible to comment on what proportion of positive PD signals at the radiocarpal joint might not have been related to synovial inflammation.

#### *Proposed MSUS definition of active disease*

Since there is no universally accepted joint set for assessment of global disease activity it follows that there is no accepted ultrasonographic definition of active disease either. Therefore, for this study, the definition of active disease, and the threshold at which DMARD therapy would be escalated, was devised pragmatically to represent a situation that balanced the sensitivity of the assessment against the possibility of false positive findings caused by PD artefact. The proposed MSUS definition defines active disease in a categorical manner (i.e. active or inactive) and is therefore more akin to the description of Minimal Disease Activity (329) than a continuous composite disease activity score. Using the available results, it was possible to determine what impact varying the PD joint count thresholds might have upon the sensitivity of the assessment and overall impression of disease activity (Tables 33 – 35). Using a PD joint count threshold of 2, 107 (26%) MSUS assessments were classified as active disease. When the PD joint count threshold was reduced to 1, 244 (59%) were classified as active disease; whereas, when the PD joint count threshold was increased to 3, only 39 assessments (9%) were classified as active disease. These changes were evident whether the indication for MSUS assessment was DAS28 LDAS or moderate disease activity with minimal clinical synovitis. Clearly, if a lowered PD joint count threshold of 1 had been used, it is likely that the MSUS assessment group as a whole would have been treated even more aggressively than they were using a threshold of 2. However, the potential of misinterpreting PD artefact in a single joint remains, and therefore such a low threshold may have lead to some participants being inappropriately exposed to the risks of intensive combination and/or DMARD therapy. By contrast, using an increased PD joint count threshold of 3 clearly restricts the sensitivity of the MSUS assessment. Whilst the stricter definition of active disease further reduces the risk of misclassifying PD artefact, it is quite unlikely that so few treatment decisions being altered would make any benefits on clinical and radiological outcomes too small to detect. It is worth noting that these assumptions on the impact of different PD joint count thresholds on treatment decisions were based upon data which was collected when the threshold was always 2. The MSUS findings observed during any given consultation will have been directly influenced by the treatment decisions made during preceding consultations (i.e. a single treatment escalation decision impacts upon all future disease activity assessments). Therefore, it is quite possible that the different PD joint count thresholds would have lead to different sequences of DMARD escalation and different sequences of subsequent clinical and MSUS disease activity assessments. Thus, separate longitudinal studies may be required to fully understand the impact of different PD joint count thresholds on DMARD escalation decisions and clinical and radiological outcomes.

### **5.8.3 Workload implications of regularly performing MSUS disease activity assessment**

By extrapolating the findings of this study to routine care, approximately 55% of all early RA clinic consultations would also need to accommodate MSUS assessment of disease activity. If the

average outpatient consultation takes between 10-20 minutes (335) adding MSUS examination to existing clinical practice will substantially prolong the consultation time required for a significant number of patients and may actually lead to a reduction in the number of patients it is possible to review in an already busy clinic. Therefore, if MSUS examination were to be used routinely alternative assessment models may need to be considered. In particular, if it is impractical for the clinician to conduct the examination during the routine appointment, reconfiguring how outpatient clinics operate may be required (see below)

If patients do achieve a stable low disease activity state that is no longer associated with on-going ultrasonographic evidence of active synovitis, it may be possible to use several consecutive MSUS assessments to confirm attainment of a stable disease state and then only consider repeating the MSUS assessment if there is a change in the patient's condition. In this way, it might be possible to still achieve the potential benefits of disease activity monitoring using MSUS (in particular the unexpected discovery of active subclinical synovitis) without having to perform multiple, negative assessments that are no longer contributing directly to clinical care. From the available results it is clear that a subset of participants underwent multiple MSUS assessments with relatively low rates of active synovitis being identified (Graph 9). When participants had undergone a lesser number of MSUS assessments the frequency that active disease was identified was notably higher (Table 36). Thus, by extension, as the number of MSUS assessments performed on a single patient increases, the probability of identifying activity synovitis decreases. Indeed, the risk of identifying active synovitis following 4 or 5 consecutive negative assessments was significantly less than after one negative assessment. Therefore, the results suggest that if patients have undergone 4 or 5 consecutive negative MSUS assessments the risk of identifying synovitis during future assessments becomes too small to justify the time and cost needed to perform the examination. Since there is still a small risk of identifying synovitis it may be prudent to perform a final MSUS assessment after a further 3-6 months to determine whether there has been any recurrence of active disease in the intervening period

<b>Number of MSUS assessments per participant</b>	<b>Number of participants</b>	<b>Total number of MSUS assessments</b>	<b>Mean rate (percentage) of positive MSUS assessments</b>
1 – 5	11	32	66%
6 – 10	29	240	32%
11 – 15	13	160	9%

**Table 36 – Rate of positive MSUS assessments in relation to number of MSUS assessments performed per patient**

When the number of MSUS assessments that identified active synovitis (i.e the positive assessments) were compared to the number of MSUS assessments for each patient it also became apparent that repeatedly performing MSUS assessments on some patients was not contributing additional disease activity information. For participants who'd undergone between 1 and 9 MSUS assessments the proportion of positive assessments appeared quite varied. In contrast, the proportion of positive MSUS assessments was substantially lower if participants had undergone 10 or more assessments. Interestingly, these repeated assessments were still

identifying evidence of synovial hypertrophy (i.e. evidence of previous joint involvement); however, it was frequently not associated with sufficient PD signal to meet the definition of.

### **Proposals for integration of MSUS assessment into routine practice**

- MSUS assessment should occur of the same day as the clinical assessment to provide contemporaneous assessments of disease activity. The requirement for MSUS assessment should be decided following clinical assessment of disease activity
- MSUS assessments can be performed by any appropriately experienced sonographer; e.g. rheumatologist, specialist nurse, extended scope physiotherapist, occupational therapist, sonographer
- Indication for MSUS assessment should be restricted to occasions when MSUS findings will change both impression of disease activity and treatment decisions: eg –
  - i. moderate disease activity but minimal clinical synovitis
  - ii. DAS28 LDAS and/or remission – less clear, MSUS assessment has not been associated with superior clinical outcomes
- MSUS joint set should contain sufficient joints to represent global disease activity, Infrequently affected or highly complex joints should be excluded to reduce scanning time. This approach could include adaptive scanning of additional symptomatic joints when necessary
- MSUS definition of active disease should be sufficiently sensitive to allow impact on DMARD escalation to positively impact upon outcomes. Not so far demonstrated by this research
- Grading of synovial hypertrophy findings is not necessary. PD findings to be graded semi-quantitatively using standardised scores (e.g. Szkudlarek et al (288) or Hammer et al(289))
- Proposed MSUS joint sets and/or definitions of active disease may eventually be superseded by findings of OMERACT Imaging Working Group
- Stop further MSUS assessment after 4 consecutive negative assessments; applied to this study, this would have lead to 98 fewer (23%) MSUS assessments being performed

#### **5.8.4 Impact of MSUS assessment upon DMARD escalation decisions**

Despite repeated evidence that current definitions of RA remission do not necessarily equate to absence of ultrasonographic evidence of synovitis (121,200,336,337), there have not yet been any studies that describe how MSUS findings alter treatment decisions or outcomes, as part of a treat-to-target strategy in early RA. Similarly, there have not yet been any studies published that report the rate of ultrasonographic synovitis in RA patients with either moderate disease activity and minimal clinical synovitis and/or those in whom clinical disease activity measures may have been elevated by external influences (such as fibromyalgic-RA overlap).

Altogether, 292 (71%) paired sets of DAS28 and MSUS assessments 'agreed' on the overall disease activity assessment and would not have lead to substantially different DMARD escalation decisions (Table 33). However, pooling the results for different MSUS indications may be incorrect since they represent quite different clinical scenarios where the outcomes of the MSUS

assessment had opposing influences on DMARD treatment decisions. Therefore, for clarity of discussion, the following sections will address the significance of findings for each MSUS indication separately:

#### *MSUS assessment during DAS28 LDAS*

The majority (75%) of MSUS disease activity assessments conducted on patients with low clinical disease activity did not identify evidence of active disease (Table 33). Importantly, this also implies that a sizable minority (25%) of MSUS assessments DID identify evidence of active synovitis and supported additional DMARD escalation. Interestingly, similar levels of ultrasonographically defined disease activity were identified when the instances of DAS28 LDAS were subdivided into subgroups for remission and LDAS without remission. The present results relate to all instances of LDAS with no adjustment having been made for the small subset of participants who continued to exhibit 2 or more clinically swollen joints. In time, an additional comparison of the MSUS findings in instances of LDAS stratified by the swollen joint count (SJC <2 vs SJC ≥2, per Aletaha et al (192)) will be conducted to determine whether there is a significant difference in the rates of ultrasonographic synovitis identified between those with and without clinical synovitis

Whenever there was discordance evident between DAS28 LDAS and the MSUS disease activity assessment, participants were considered for further intensification of their DMARD therapy in an attempt to suppress the last vestiges of subclinical synovitis. This would have led to treatment changes at levels of clinical disease activity that in the DAS28 assessment group, and usual care, would be considered too low to justify further DMARD escalation. Hence, it is likely that, for the duration of the research the MSUS assessment group as a whole received more intensive therapy (including earlier escalation and higher rates of combination therapy and/or biologic therapy) than the DAS28 group. In most of these instances, patients reported relatively low levels of symptoms (mean global health 10cm VAS 11.27, mean pain 10cm VAS 11.84) and it does not appear that increasingly intense therapy was associated with significant improvements in outcomes. Increasingly aggressive combinations of DMARDs and/or biologic therapies are potentially associated with an increased risk of toxicity overall (144) and infection particularly (338). Therefore, it is possible that relentlessly escalating DMARD therapy in patients who have very low symptom burdens may become unacceptable because of: i. increasingly complicated treatment regimens with low symptomatic benefit: treatment burden ratios and/or ii. an unfavourable benefit: risk ratio. At present, the data relating to the rate and type of adverse events in either the MSUS or DAS28 group is not available for analysis. In due course, once the relevant data is available, a careful between group comparison of the number, type and severity of adverse events occurring in both assessment groups is planned in order to determine the impact of MSUS-driven DMARD escalation on adverse event rates

Previously published studies have predominantly reported the incidence of MSUS findings in RA patients with remission and not LDAS (121,217,336,337,339). Direct comparisons between those studies and these results is limited because none of the studies have specifically reported the prevalence of PD signal in DAS28 LDAS and because there are notable methodological

variations, relating to differences in joint set and how PD findings are reported. Nevertheless, the consensus appears to be that a significant proportion of patients continue to exhibit some degree of PD signal, even when they fulfil existing remission criteria. Based upon published results, the proportion of RA patients in remission who continue to exhibit PD signal in at least one joint lies between 40% (217) and 62% (339). Similarly, this study found that 60% of patients in DAS28 remission, and 59% in DAS28 LDAS, exhibited evidence of PD signal in at least one joint (Table 33). However, whilst these rates are similar it is worth restating that the previously published studies used different (often subjective) definitions of remission and different MSUS joint sets.

*MSUS assessment during DAS28 moderate disease activity but minimal clinical synovitis*

This study has shown that the majority of MSUS assessments conducted during instances of DAS28 moderate disease activity but minimal clinical synovitis did not identify active synovitis (Table 32). Fifty six percent of MSUS assessments identified PD signal in at least one joint, however, this fell to 33% when defining active disease as PD signal in at least 2 joints. MSUS assessment identified a higher rate of active synovitis during instances of DAS28 moderate disease activity but minimal clinical synovitis compared to DAS28 LDAS (33% vs 25%) but this did not reach statistical significance (Fishers exact test;  $p=0.2782$ ). It is perhaps surprising that there was not a higher incidence of ultrasonographic synovitis identified during instances of DAS28 moderate disease activity but minimal clinical synovitis, since comparisons with instances of DAS28 LDAS had shown the individual clinical disease activity variables (excepting swollen joint count) to be notably higher (Table 29)

Most published research has focussed on the ultrasonographic findings in RA patients in clinical remission. There are very few published results against which the DAS28 moderate disease activity findings from this study can be compared. A single study by Brown et al did describe the ultrasonographic and MRI findings of a subset of 24 RA patients who qualified as DAS28 moderate disease activity (121). However, this cohort is biased by selection since the referring clinician subjectively considered the patient to be in clinical remission and moderate disease activity was only diagnosed once the DAS28 assessment had been applied retrospectively. MSUS assessment of the dominant wrist and MCP joints identified synovial hypertrophy in 23 patients (95.8%) and PD signal in 20 patients (83.3%). Clearly, these rates of PD signal are somewhat higher than those identified by this research which used a broader MSUS joint set but only identified PD signal in the joints of 56% of patients with moderate disease activity but minimal clinical synovitis. The participants examined by this research are biased by indication since MSUS examination was only performed for the subset of instances of moderate disease activity when there was 1 or no clinically swollen joints. By contrast, Brown et al performed MSUS assessment on all patients with moderate disease activity. Furthermore, the clinical features of the cohort examined by Brown et al do suggest a higher degree of disease activity overall. Whilst, the mean group DAS28 score is not specifically reported, the mean swollen joint count was 4 (versus 0.42 for this research) suggesting that the Brown cohort exhibited a higher rate of clinical synovitis that could also have been associated with an increased rate of underlying MSUS abnormalities.

Despite the realisation that other joint pathologies may contribute to elevated RA disease activity assessments there have been few studies that describe the ultrasonographic findings in RA-overlap syndromes. In fibromyalgia, most current imaging research is focussed on identifying whether or not there are specific functional MRI changes present that may be associated with abnormal central pain-related brain activity (340). A single study has previously compared PD findings at peripheral enthesal points between 30 patients with psoriatic arthritis and 30 with fibromyalgia (341). Any enthesal PD signal abnormalities was identified in a significantly higher proportion of psoriatic arthritis patients (100% vs 80%,  $p=0.01$ ), who were also more likely to exhibit 'inflammatory changes' (70% vs 23%,  $p=0.001$ ). Whilst fibromyalgia patients exhibited significantly lower rates of inflammatory enthesal changes, it is interesting that there was still some PD signal abnormalities identified and such a high rate of enthesal abnormalities. Unfortunately, it is not possible to directly compare the results of Marchesoni et al to the findings of this study, since different joint recesses were examined for different sonographic abnormalities

Compared to fibromyalgia the presence of ultrasonographic synovial abnormalities has been studied more extensively in osteoarthritis. The emerging consensus suggests that osteoarthritis can also be associated with findings of joint effusion, synovial hypertrophy and PD signal (212,342,343). In the UK population the prevalence of symptomatic hand osteoarthritis is estimated between 12-30% (344,345) and approximately 4.4 million people are estimated to have radiographic hand osteoarthritis (Arthritis Research Campaign, 2002). Therefore, it is highly likely that a proportion of RA patients will also develop degenerative joint disease. Since joint osteoarthritis may also be associated with inflammatory like synovial lesions on MSUS, the specificity of disease activity assessment by MSUS examination in patients with RA-osteoarthritis overlap is potentially reduced. A number of participants in this study exhibited clinical (e.g. nodular PiPj and DiPj changes) and radiological features of osteoarthritis at presentation. It is quite possible that, during MSUS disease activity assessments PD signal related to osteoarthritis was attributed to RA synovitis instead and contributed to the RA disease activity state being classified as active. Unfortunately, it is not possible to estimate the potential impact of co-existing osteoarthritis on MSUS disease activity assessment findings since the rate of osteoarthritis at baseline was not specifically recorded. Further, the MSUS assessment focussed predominantly on the presence of synovial hypertrophy and (especially) PD signal. The presence of other joint abnormalities (e.g. osteophytes), that may allow clearer discrimination between inflammatory and degenerative synovial findings, were not recorded. Until recently, treatment of osteoarthritis focused on symptom relieving measures and medications. The increasing realisation that some sub-types of osteoarthritis are associated with the presence of inflammatory-like lesions has lead some investigators to trial treatment with immunomodulatory agents. So far, clinical trials have suggested that methotrexate (346), hydroxychloroquine (347), anti-TNF $\alpha$  blocking therapies (348) and systemic corticosteroids (349) can all have either symptom relieving (methotrexate, hydroxychloroquine, intra-muscular methylprednisolone) and/or radiographic progression limiting (anti-TNF $\alpha$  blocking agents) beneficial effects. Thus, even if participant's DMARD therapy was escalated because inflammatory osteoarthritic lesions were identified during MSUS assessment it is possible that participants may still have experienced some symptomatic benefit from the therapy change. However, disentangling whether any benefit was related to improvements in

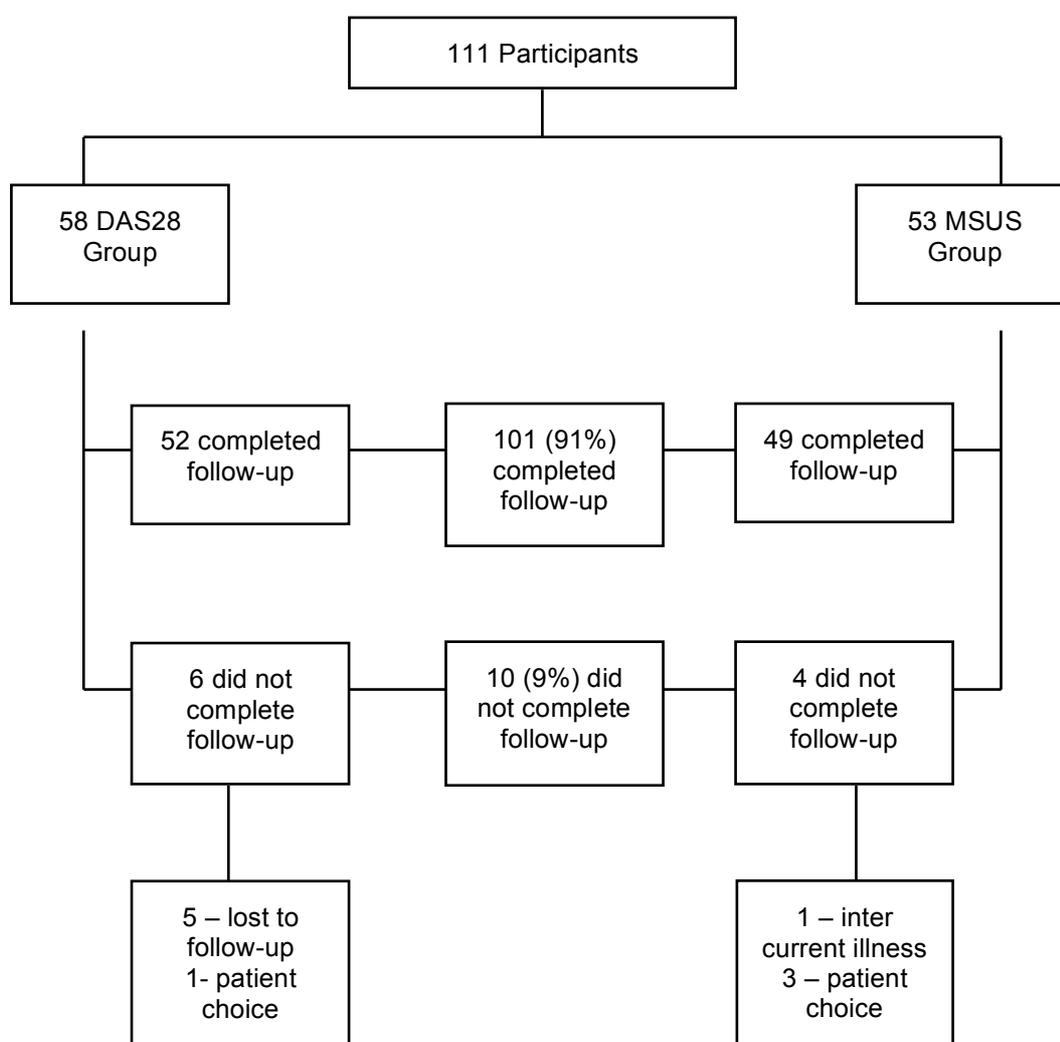
either RA and/or osteoarthritic inflammatory activity will not be possible using the currently available results.

## **6. Impact of MSUS Disease Activity Assessment upon Clinical Outcomes**

During the process of writing this thesis the period of follow-up for the research participants was completed, therefore, the following chapter shall present the results of an analysis of the clinical outcomes for the DAS28 and MSUS assessment groups. Due to the timing of the submission deadline it will not be possible to present the results of any of the radiological outcomes nor the adverse event rates. These data will be presented in the academic press in due course and will be deposited with the thesis in University records for future reference.

## 6.1 Clarification of disposition of study participants

Fifty-eight participants were randomised to the DAS28 (control) assessment group and 53 were randomised to the MSUS assessment group. By the completion of the research, 52 participants remained within the DAS28 group (10% dropout) and 49 participants remained in the MSUS group (8% drop-out). The overall drop out rate was 9% and there was no difference in the proportion of drop-outs from either group ( $p = 0.74$ , Fishers exact test). Based on these findings a continuation of the preliminary consort diagram (Figure 12) is presented below (Figure 13).



**Figure 13** – Continuation of consort diagram (Figure 13) to include disposition of all research participants between recruitment and final study visit

## 6.2 Comment on the statistical analysis

The following analysis is restricted to the ACR core set variables (i.e. the clinical outcomes). All analyses have been conducted using Graphpad Prism Version 6 ([www.graphpad.com](http://www.graphpad.com)). However, the data have not been subjected to the checking and quality control procedures of the Robertson Centre for Biostatistics, University of Glasgow that usually accompany the completion of a prospective clinical trial. In due course a final statistical analysis of the whole dataset will be conducted with the assistance of the Robertson Centre for Biostatistics statisticians prior to formal reporting of the research's main clinical and radiological outcomes.

For every ACR core set variable, frequency distribution analysis demonstrated quite significant variability in the skew and distribution at each of the 3 month assessment time points (Table 37). Values of skew greater than 1 suggest that the distribution of the data within a group is very unlikely to be symmetrical; therefore, since it appears that a significant number of the data groups are not distributed normally, non-parametric between-group comparisons have been conducted. Unless otherwise stated, a p value less than 0.05 was considered statistically significant

ACR Core Set Variable	DAS28 Group	MSUS Group
DAS44	0.56 – 1.64	0.10 – 1.24
HAQ	-0.22 – 1.17	0.18 – 1.50
44 Swollen Joint Count	0.92 – 4.40	0.21 – 4.23
Ritchie Articular Index	1.52 – 3.53	0.57 – 4.57
ESR	1.13 – 3.59	1.03 – 3.04
CRP	2.25 - 5.0	1.71 - 5.50
Global Health 10cm VAS	0.08 – 1.86	0.02 – 1.87
Pain 10cm VAS	0.38 – 1.61	0.33 – 1.70

**Table 37** – Range of skew across the whole follow-up period for each ACR core set variable

## 6.3 Overall change in ACR core set variables

Table 38 describes the overall change in each ACR core set variable between baseline and completion of 18 months follow-up. There were no statistically significant between-group differences in the mean change from baseline of any variable. In the MSUS assessment group, for each ACR core set variable, the 18 month value tended to be lower than the corresponding value for the DAS28 assessment group. Furthermore, excepting ESR, the mean change from baseline tended to be slightly higher for the MSUS assessment group.

In both assessment groups, there was a statistically significant improvement in all of the ACR core set variables measured between baseline and 18 months. Furthermore, the time period that corresponded with the fastest rate of improvement was consistently between baseline and 3 months for both groups. Equally, for most ACR core set variables, very few additional

improvements were measured between 3 months and 18 months. The specific patterns of change of each of the ACR core set variables will be discussed in more detail in the following sections 6.4 to 6.6

<b>ACR Core Set Variable</b>	<b>DAS28 Group n = 57</b>	<b>MSUS Group n = 54</b>	<b>Difference between means (95%CI)</b>	<b>p (DAS28 vs MSUS)</b>
<b>DAS44</b>				
Baseline	4.43 (3.19 – 5.42)	4.32 (3.63 – 5.25)		0.98 (MW)
18 months	1.73 (0.92 – 2.63)	1.25 (0.64 – 2.43)		0.12 (MW)
Mean change <i>mean (SD)</i>	-2.58 (1.57)	-2.69 (1.41)	-0.11 (-0.70, 0.48)	0.72 (t test)
<b>HAQ</b>				
Baseline	1.63 (0.88 – 2.13)	1.5 (1.0 – 2.0)		0.78 (MW)
18 months	0.5 (0.0 – 1.38)	0.0 (0.0 – 0.91)		0.062 (MW)
Mean change <i>mean (SD)</i>	-0.79 (0.70)	-1.02 (0.81)	-0.23 (-0.52, 0.07)	0.14 (t test)
<b>44 Swollen Joint Count</b>				
Baseline	8 (5 – 14)	9 (5 – 13)		0.97 (MW)
18 months	0 (0 – 1)	0 (0-0)		0.07 (MW)
Mean change <i>mean (SD)</i>	-8.6 (6.6)	-8.2 (5.3)	0.37 (-2.0, 1.2)	0.76 (t test)
<b>Ritchie Articular Index</b>				
Baseline	17 (10 – 28)	18 (11 – 28)		0.89 (MW)
18 months	2 (0 – 6)	0 (0 – 5)		0.23 (MW)
Mean change <i>mean (SD)</i>	-15.7 (13.8)	-14.1 (11.5)	1.59 (-3.4, 6.6)	0.53 (t test)
<b>ESR</b>				
Baseline	25 (11 – 52)	35 (16 – 43)		0.90 (MW)
18 months	15 (7 – 20)	8 (5 – 23)		0.15 (MW)
Mean change <i>mean (SD)</i>	-20 (26)	-20 (25)	0.56 (-9.4, 10.6)	0.91 (t test)
<b>CRP</b>				
Baseline	15 (7.4 – 44.5)	19.5 (5.9 – 50.5)		0.81 (MW)
18 months	3.8 (1.8 – 9.3)	4.0 (1.6 – 10.0)		0.71 (MW)
Mean change <i>mean (SD)</i>	-29.4 (60.1)	-28.8 (42.6)	0.61 (-20.0, 21.2)	0.95 (t test)

**Table 38** – Overall change in each ACR Core Set Variable between baseline and 18 months  
Unless stated values are medians (IQR)  
MW – Mann Whitney U test; t test = unpaired, 2 tailed t test

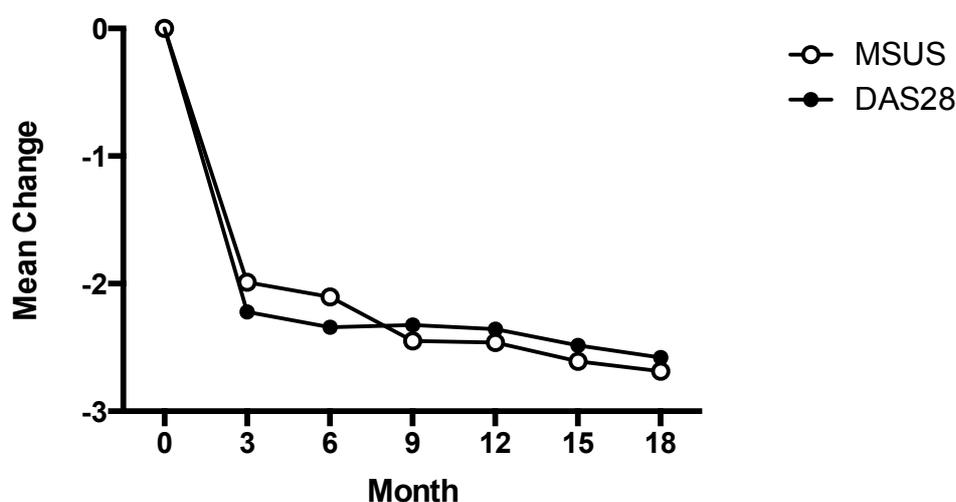
ACR Core Set Variable	DAS28 Group n = 58	MSUS Group n = 53	Difference between means (95%CI)	p (DAS28 vs MSUS)
<b>Global Health 10cm VAS</b>				
Baseline	54 (31 – 79)	54.5 (39.8 – 67.2)		0.74 (MW)
18 months	6 (1 – 29)	2.5 (0 – 22.3)		0.28 (MW)
Mean change <i>mean (SD)</i>	-40.5 (34.7)	-43.4 (29.0)	-2.9 (-15.6, 9.7)	0.65 (t test)
<b>Pain 10cm VAS</b>				
Baseline	51 (34.3 – 69.8)	48.5 (33.5 – 57.5)		0.30 (MW)
18 months	7 (1 – 19)	3.5 (0 – 23)		0.33 (MW)
Mean change <i>mean (SD)</i>	-34.9 (31.6)	-32.4 (27.0)	2.5 (-9.2, 14.2)	0.68 (t test)

**Table 38** (continued)– Overall change in each ACR Core Set Variable between baseline and 18 months.

## 6.4 Change in DAS44

### 6.4.1 Mean Change in DAS44 from baseline

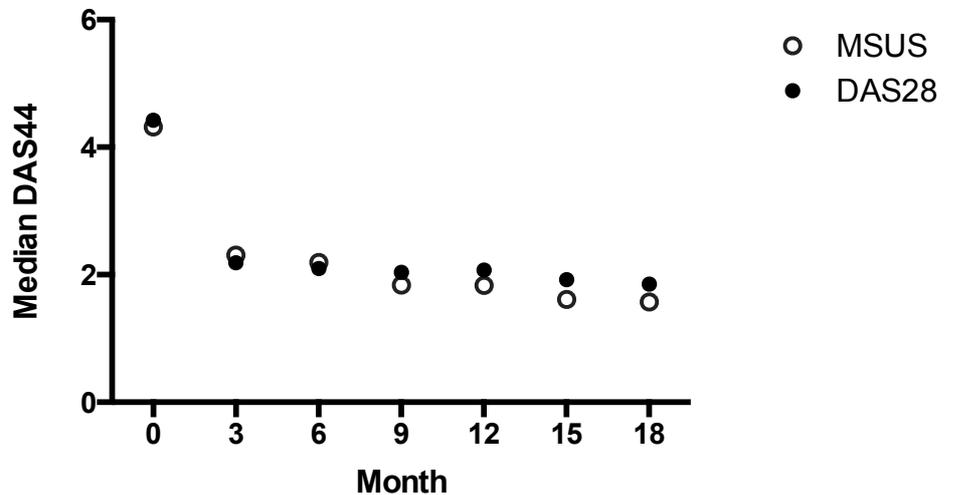
The primary clinical outcome measure was the mean change in DAS44 exhibited by both groups. The original sample size calculation was based upon detecting a between group difference of 1.1 or higher in the mean change of DAS44 after 18 months rather than a difference in the absolute DAS44 values. The overall difference in the mean change in DAS44 was 0.11, suggesting that both groups had experienced a similar improvement in DAS44. At every assessment time point, both groups exhibited a similar mean change in DAS44 from baseline (Graph 19). There were no statistically significant between-group differences in the mean change from baseline of DAS44 at any of the assessment time points (p 0.36-0.83) and the rate of change in DAS44 appeared similar for both groups.



**Graph 19** – Mean change from baseline of DAS44

### 6.4.2 DAS44 at each time point

For both assessment groups there was a significant overall improvement in DAS44 between the baseline and 18 months follow-up visits (Graph 20,  $p < 0.0001$ ). There were no statistically significant between group differences in the distribution of DAS44 values measured at any of the assessment time points ( $p = 0.18 - 0.71$ ).



Graph 20 – Median DAS44 score at each assessment time point

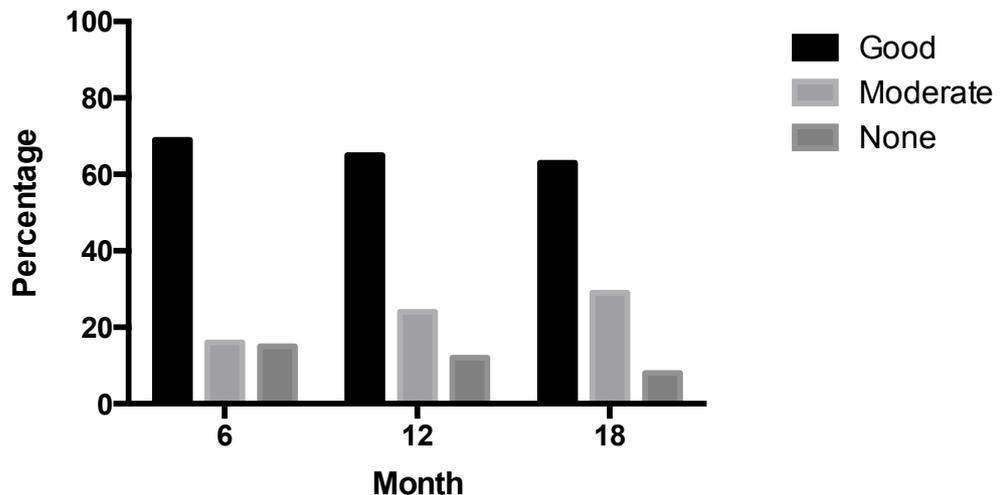
The steepest rate of change of DAS44 for both groups was measured between the baseline and 3 month assessment visits. Between follow-up months 3 and 18 there was continued improvement measured in the median DAS44 for both groups. In the DAS28 group, median DAS44 fell from 1.79 (IQR 1.04 – 2.95) to 1.73 (IQR 0.93 – 2.61) ( $p = 0.018$ ). In the MSUS group, median DAS44 fell from 2.37 (IQR 1.22 – 3.12) to 1.25 (IQR 0.64 – 2.43) ( $p < 0.0001$ ). There was no statistically significant difference in the mean change in DAS44 between either group between follow-up months 3 and 18 (mean change DAS44: DAS28 group -0.39 vs MSUS group -0.72,  $p = 0.19$ ).

Overall, both groups appear to have been exposed to a similar cumulative inflammatory burden. Using each patient's mean DAS44 between baseline and month 18 as a surrogate for inflammatory exposure, there was no significant difference identified between the medians measured in either group (DAS28 2.23 (IQR 1.55 – 3.18) vs MSUS 2.27 (IQR 1.38 – 2.85),  $p = 0.64$ ).

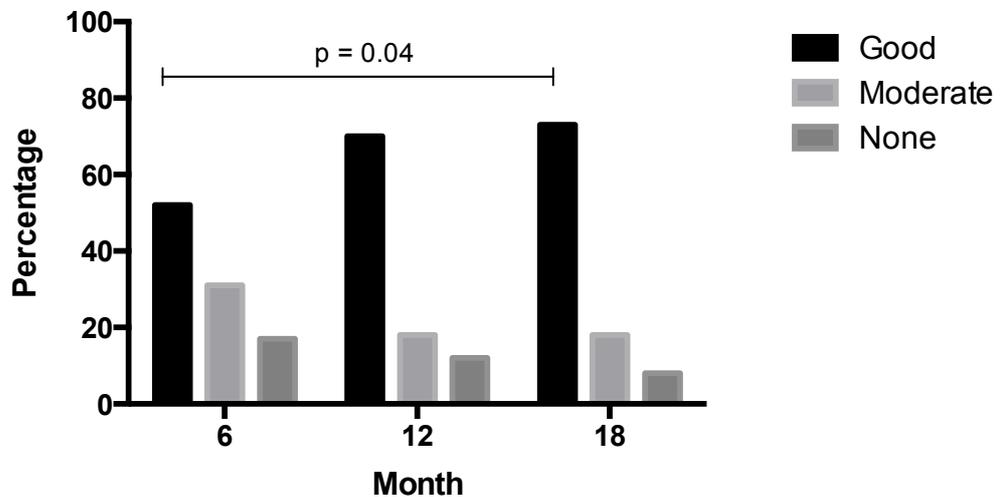
### 6.4.3 EULAR response rates

Using both the final value and the absolute change from baseline of the DAS44, the EULAR response rates after 6, 12 and 18 months follow-up were calculated for each group (297) (Graphs 21 and 22). At each 6 monthly assessment time point between group comparisons were conducted (Fisher's exact test) of the proportion of participants in each group meeting each of the different EULAR response definitions. Participants were excluded from the comparison if there was no DAS44 data available for the follow-up visits

At each time point, there was no significant difference between the proportion of participants in either group fulfilling any of the EULAR response definitions. Over the follow-up period there was an upward trend in the proportion of MSUS group participants who achieved a EULAR good response (6 months 52%; 12 months 70%; 18 months 73%) which was statistically significant between months 6 and 18 ( $p = 0.039$ ). The proportion of DAS28 groups who achieved a EULAR good response remained relatively static over the follow-up period (6 months 69%; 12 months 65%; 18 months 63%)



Graph 21 – DAS28 Group – percentage of participants fulfilling EULAR response criteria



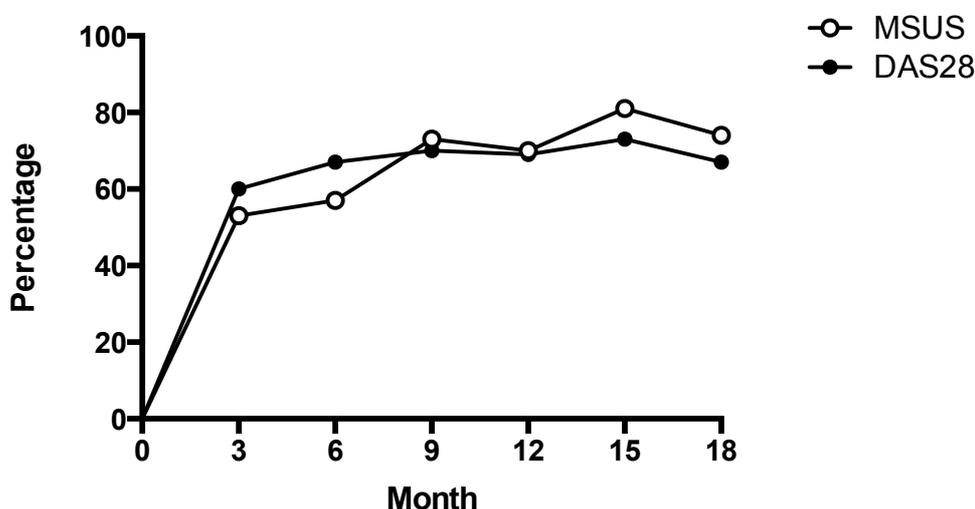
Graph 22 - MSUS Group – percentage of participants fulfilling EULAR response criteria

#### 6.4.4 Rate of Low Disease Activity and Remission

The DAS44 results were analysed to determine what proportion of participants within each group met DAS44 LDAS (DAS44 <2.4) and remission (DAS44 <1.6) definitions at any point during the follow-up period and what proportion achieved remission on one or more occasions (e.g. persistent remission).

##### *Rates of DAS44 Low Disease Activity (LDAS)*

In this section, DAS44 LDAS includes the subset of participants who had also attained DAS44 remission. After 18 months of treatment, a similar proportion of participants within both groups had achieved DAS44 LDAS (DAS28 67% vs MSUS 74%,  $p$  0.51) (Graph 23). There was no statistically significant between group differences in the number of participants in either group achieving DAS44 LDAS at any of the assessment time points ( $p$  0.32 – 1.0). Indeed, between follow-up months 12 until 18 the number and proportion of participants within each group achieving DAS44 was virtually identical.



**Graph 23** - Percentage of participants within each group achieving DAS44 LDAS (DAS44 < 2.4) between follow-up months 3 to 18

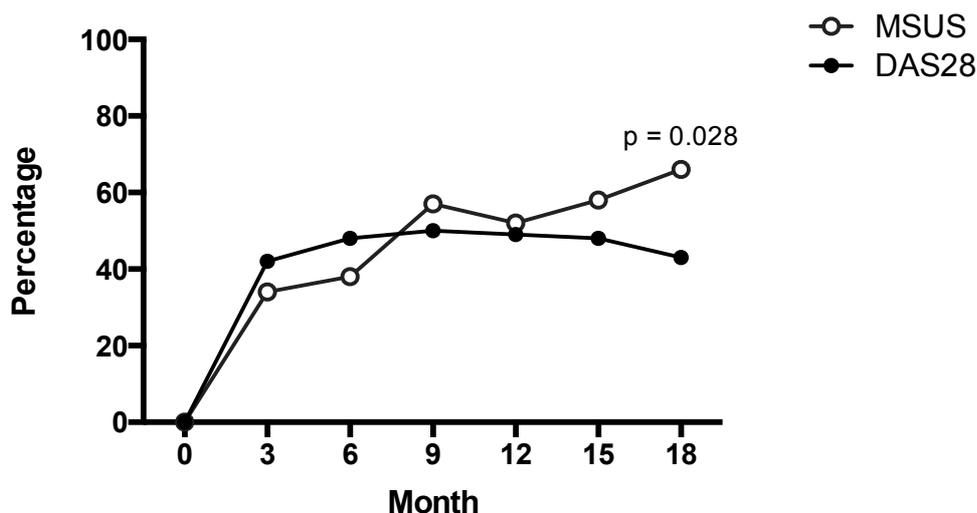
Within the DAS28 assessment group, the number of participants achieving DAS44 LDAS remained relatively static between follow-up months 3 and 18 ( $n$  = 33-38). By contrast, there was an incremental increase in the number of participants within the MSUS assessment group who achieved DAS44 LDAS. In fact, there was a statistically significant increase in the number of MSUS participants who achieved DAS44 LDAS between follow-up months 3 and 9 (52% to 73%,  $p$  0.003), months 3 and 15 (52% to 81%,  $p$  0.0032) and months 3 and 18 (52% to 74%,  $p$  0.04).

##### *Rates of DAS44 Remission*

Excepting follow-up month 18, a very similar pattern of change to that described for LDAS was observed for the proportion of participants in each group who achieved DAS44 remission (Graph 24). After 18 months, a significantly higher number of patients within the MSUS group had achieved DAS44 remission compared to the DAS28 group (66% vs 43%,  $p$  0.028). Prior to month

18, there were no significant differences observed in the proportion of participants achieving DAS44 remission in either group ( $p = 0.32 - 0.84$ ).

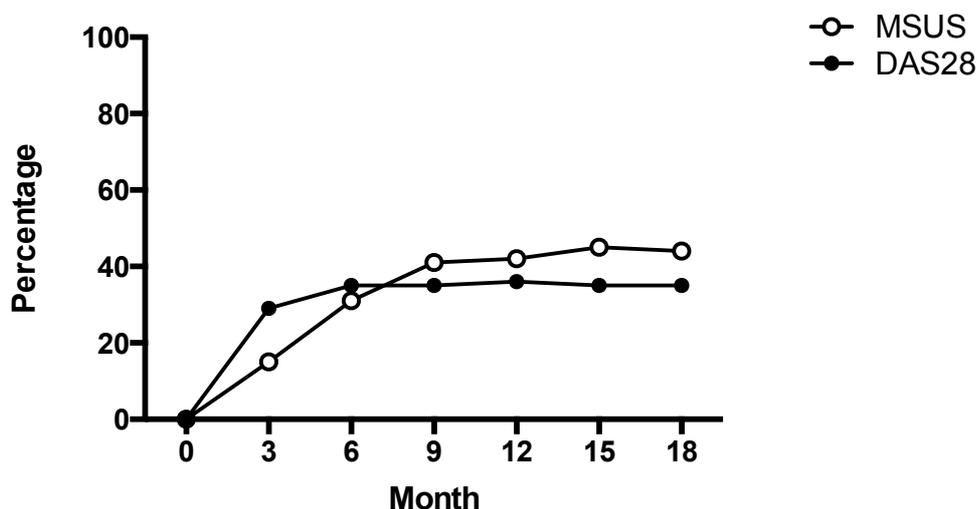
In the MSUS group, the proportion of participants who attained DAS44 remission increased over the follow-up period. A statistically significant increase was observed between follow-up months 3 and 9 (34% to 57%,  $p = 0.030$ ), months 3 and 15 (33% to 58%,  $p = 0.017$ ), months 3 and 18 (33% to 66%,  $p = 0.0016$ ) months 6 and 15 (38% to 58%,  $p = 0.047$ ) and months 6 and 18 (38% to 66%,  $p = 0.0057$ ). By contrast, there was no increase in the proportion of DAS28 group participants who attained DAS44 remission between months 3 and 18.



**Graph 24** – Percentage of participants within each group achieving DAS44 remission (DAS44 <2.4)

#### *Rates of ACR-EULAR Boolean Remission*

A very stringent definition of remission has been proposed by the ACR and EULAR based on all the common components of composite disease activity measures being virtually normal ( $TJC \leq 1$ ,  $SJC \leq 1$ , Global Health VAS  $\leq 1$ cm,  $CRP \leq 1$ mg/dl)(186). This Boolean definition of remission is perhaps a more appropriate outcome since it seems to much more closely represent the virtual absence of clinically detectable synovitis. Applying the ACR/EULAR Boolean definition of remission to the clinical outcomes of this study does attenuate the apparent impact of MSUS assessment on the DAS44 remission rates of the MSUS group (Graph 25). For both groups, the pattern of attainment over the time course is unchanged; the DAS28 group ACR/EULAR remission rate remains relatively static between months 3 and 18, whilst for the MSUS group the remission rate increases gradually. However, the proportion of participants within each group achieving ACR/EULAR remission is consistently less in both groups than the proportion who attained DAS44 remission. Further, whilst the ACR/EULAR remission rates for the MSUS group once again appear numerically higher than the DAS28 group between follow-up months 9 and 18 there is no statistically significant between group difference in the proportion of participants achieving ACR/EULAR remission at any of the time points ( $p = 0.11 - 0.84$ )



**Graph 25** – Percentage of participants attaining ACR/EULAR remission follow-up period  
(\* DAS28 vs MSUSp<0.05)

*Rates of sustained DAS44 remission*

Forty-four participants (76%) within the DAS28 group and 41 participants (77%) within the MSUS group attained DAS44 remission on at least one occasion. Further, participants within either group were equally likely to achieve DAS44 remission on at least one occasion (OR 0.9, 95%CI 0.4, 2.2).

Since the persistence of a remission state may be more suggestive of inactive disease than an assessment relating to a single time point, the odds of participants within either group attaining DAS44 remission on more than one occasion (Table 39) and on more than one consecutive occasion (Table 40) were calculated. There was no statistically significant differences in the likelihood of participants in either group attaining DAS44 remission on multiple separate occasions or consecutive occasions. Though the odds suggested participants within the MSUS group were slightly more likely to have achieved DAS44 remission on 5 or more separate occasions and 5 or more consecutive occasions.

Number of occasions of remission (not consecutive)	Odds Ratio (95% CI)	p (Fisher's exact test)
1 or more	0.9 (0.4, 2.2)	1
2 or more	1.1 (0.5, 2.3)	1
3 or more	1.4 (0.7, 2.9)	0.45
4 or more	1.1 (0.5, 2.3)	1
5 or more	0.7 (0.3, 1.7)	0.5
6	0.6 (0.2, 1.6)	0.47

**Table 39** - Odds of participants within DAS28 group attaining DAS44 remission on separate occasions compared to MSUS group

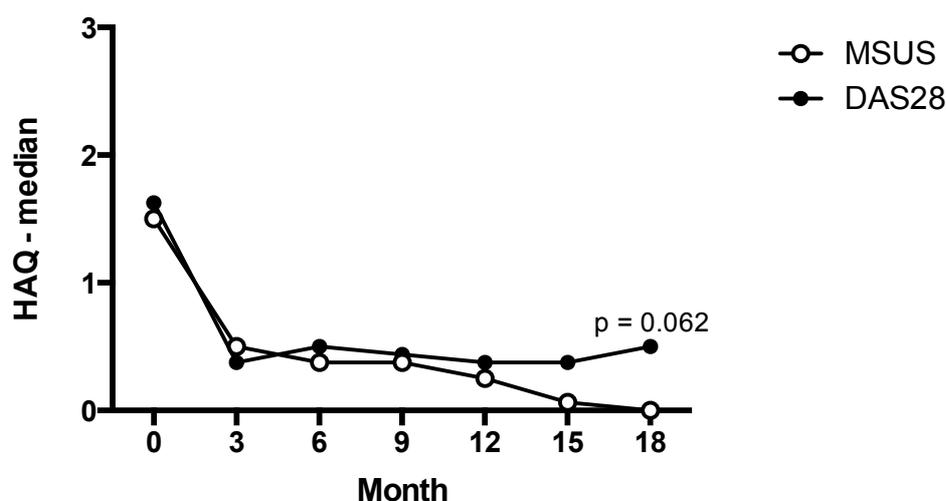
Number of consecutive occasions of remission	Odds Ratio (95% CI)	p (Fisher's exact test)
1 or more	0.6 (0.2, 1.5)	0.36
2 or more	1.4 (0.6, 2.9)	0.45
3 or more	1.2 (0.6, 2.6)	0.70
4 or more	0.9 (0.4, 2.1)	1
5 or more	0.7 (0.3, 1.8)	0.51
6	0.6 (0.2, 1.6)	0.47

**Table 40** –Odds of participants within DAS28 group attaining DAS44 remission on consecutive occasions compared to MSUS group

## 6.5 Change in HAQ

At baseline, very similar levels of functional impairment were measured in both groups by the Health Assessment Questionnaire. Both groups experienced a significant improvement in functional ability between baseline and 18 months (Graph 26). In the DAS28 group median HAQ fell from 1.63 to 0.5 ( $p < 0.0001$ , Wilcoxon matched-pairs signed rank test) and in the MSUS group median HAQ fell from 1.50 to 0.0 ( $p < 0.0001$ , Wilcoxon matched-pairs signed rank test). Eighteen month median HAQ levels were numerically lower in the MSUS group though the between group difference only reached borderline statistical significance ( $p = 0.062$ , Mann Whitney U test).

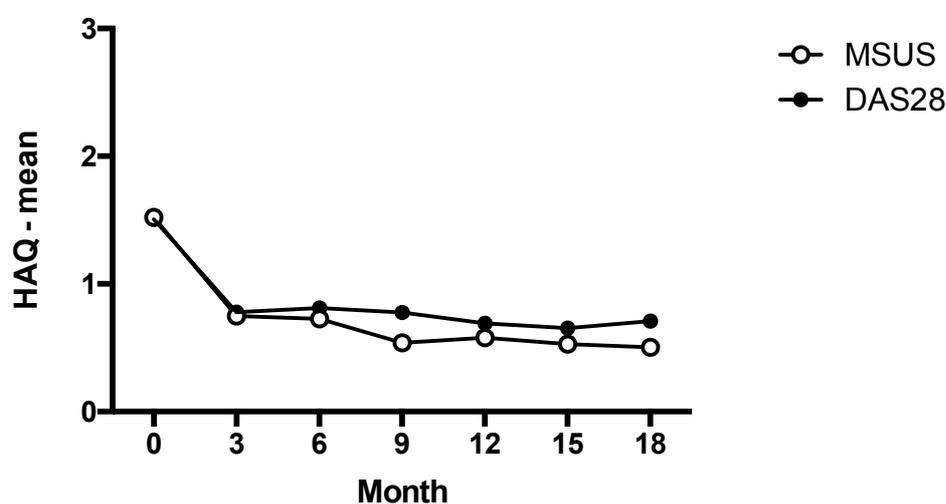
Patients in the MSUS group tended to experience a greater overall mean change in the HAQ (DAS28 -0.79 vs MSUS -1.02). The difference of the mean change from baseline of HAQ (-0.23) was not statistically significant ( $p = 0.14$ ) but, interestingly, was greater than the minimum clinically important difference (0.22).



**Graph 26** – Median HAQ at each assessment time point

Measures of skew suggested that the distribution of HAQ data for each group at each assessment time point was highly asymmetrical (range -0.22 – 1.499); hence, non-parametric

statistical tests were used to compare between group differences at each time point. The median HAQ values for the MSUS group at months 12, 15 and 18 were very low (0.25, 0.063 and 0.0 respectively) suggesting that at least 50% of the group reported a HAQ that was essentially normal. Indeed, the difference in median HAQ between the groups at months 15 (0.38 vs 0.063) and month 18 (0.5 vs 0.0) is greater than the minimal clinically important difference (300). However, reporting median HAQ values alone may under-estimate and mis-represent the level of functional impairment experienced by the upper half of the MSUS group. Calculating the mean HAQ for each group at each assessment time point reveals a slightly different pattern (Graph 27). The pattern of improvement in HAQ remains similar and both groups continue to demonstrate significant overall improvements. However, in the later follow-up months, the mean HAQ (0.52) for the MSUS group is slightly higher than indicated by the median value (0.0) suggesting that at least some of the MSUS group were still experiencing low level function impairment. Interestingly, there is virtually no between group difference in the mean HAQ value, or the rate of improvement, at any of the assessment time points.



Graph 27 - Mean HAQ at each assessment time point

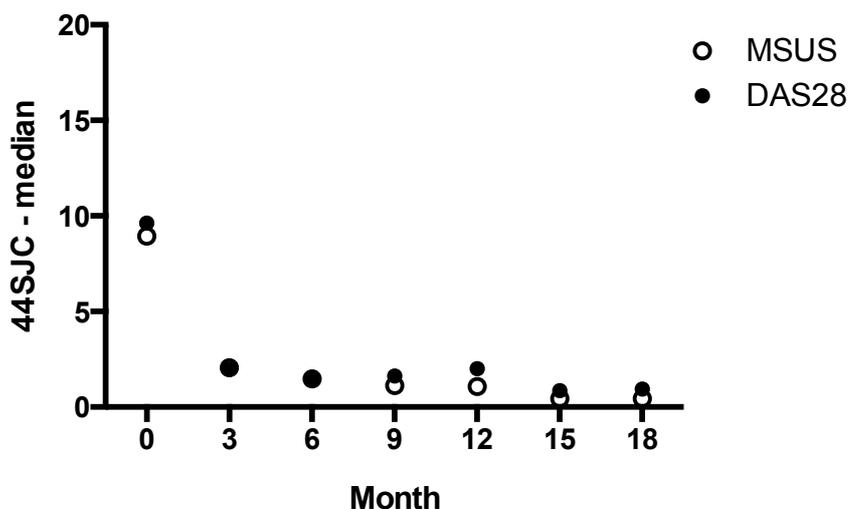
## 6.6 Change in Other ACR Core Set Variables

Both groups exhibited significant improvements from baseline of all other ACR core set variables with the steepest change evident between baseline and follow-up month 3. For each variable, between group comparisons were conducted at each time point, though none of these comparisons suggested a statistically significant difference. Similarly, none of the between group comparisons for mean change from baseline of a variable were statistically significant either. Graphs 28 to 33 depict the change from baseline of each variable by group

### 6.6.1 44 Swollen Joint Count

At the baseline assessment, both groups exhibited a similar median number of clinically swollen joints (8 vs 9, p 0.97) which fell, and remained at a very low level (0-1) from month 9 onwards for both groups (Graph 28). There were no significant between group differences in the median

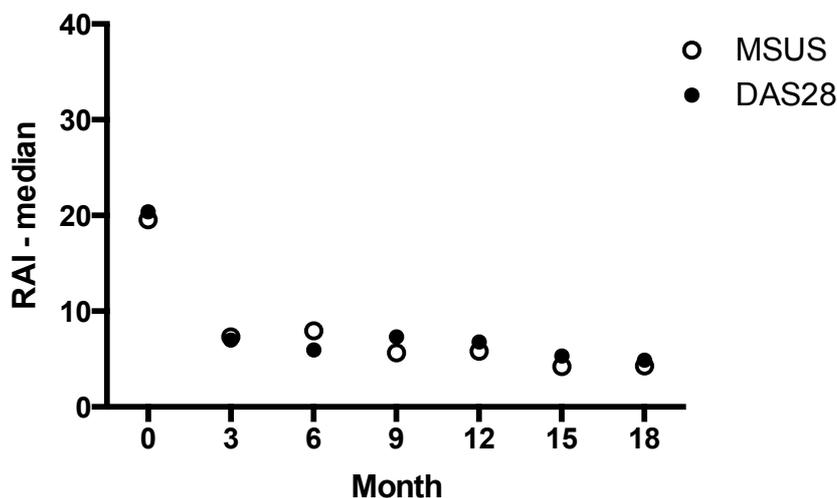
44SJC at any time point ( $p = 0.1 - 0.91$ ). Nor was there a significant differences in the mean change in 44SJC observed between either group (DAS28 -8.6 vs MSUS -8.2,  $p 0.76$ ).



Graph 28 – Median 44SJC at each assessment time point

### 6.6.2 Change in Ritchie Articular Index

There were no statistically significant differences in the RAI identified between either group at any of the assessment time points ( $p 0.29-0.87$ ) (Graph 29). At baseline, median RAI for the DAS28 group was 17 and had fallen to 2.0 after 18 months follow-up ( $p < 0.0001$ ). For the MSUS group, median baseline RAI was 18.0 and fell to 0.0 after 18 months ( $p < 0.0001$ ). There was no statistically significant difference in the mean change of RAI between either group (DAS28 -15.7 vs MSUS -14.1,  $p=0.53$ ).

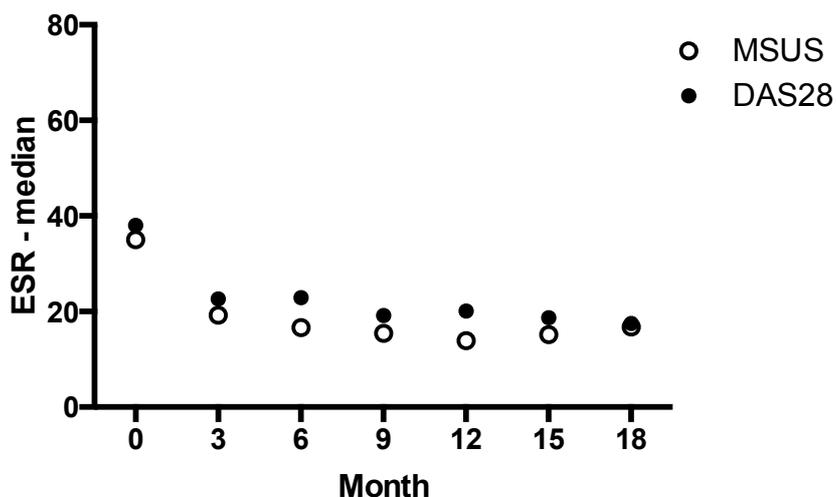


Graph 29 – Median Ritchie Articular Index at each assessment time point

### 6.6.3 Change in ESR

There were no significant difference identified in the mean overall change in ESR between baseline and 18 months in either group ( $p 0.91$ ) (Graph 30). There were no statistically significant differences in the ESR values between the groups at any of the assessment time points ( $p 0.067-0.85$ ). Median ESR values at 18 months for both groups were almost within the normal range

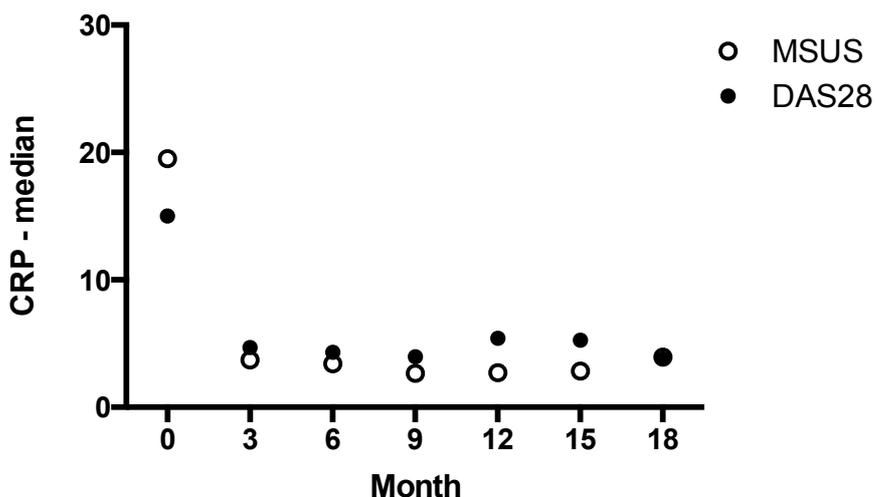
(DAS28 15 (IQR 7– 20) vs MSUS 8 (IQR 5-23)). However, the interquartile ranges suggest that the distribution of values was broad.



Graph 30 – Median ESR at each assessment time point

#### 6.6.4 Change in CRP

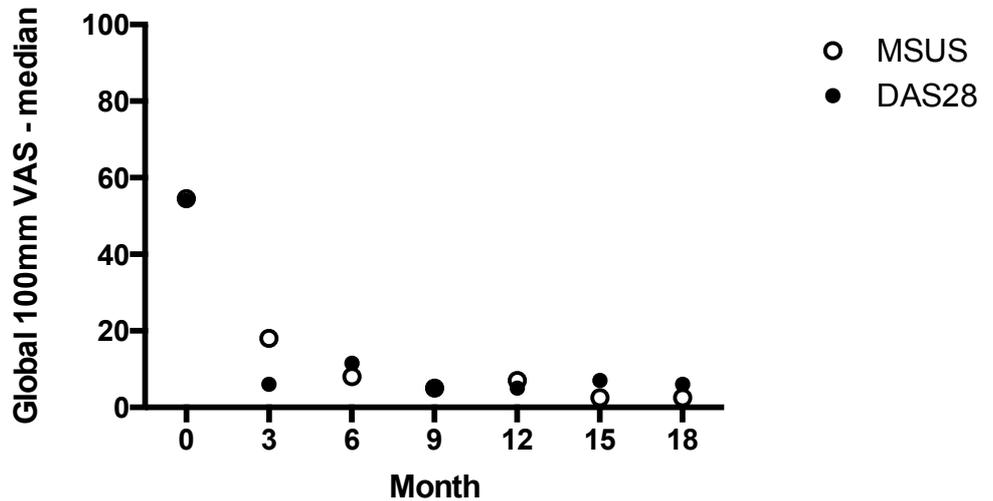
At a group level, participants in both groups exhibited moderate elevation of CRP at baseline (median CRP: DAS28 15, MSUS 19.5;  $p$  0.81) (Graph 31); however, the interquartile ranges were heavily negatively skewed (IQR: DAS28 7.4-44.5; MSUS 5.9-50.5) and there were notable outliers in both groups (maximum baseline CRP: DAS28 314; MSUS 191). There were no significant between group differences in the CRP values at any time point ( $p=$  0.08-0.81) nor in the mean change from baseline ( $p =$  0.95). By follow-up month 18, the range of the CRP values for each group appeared to have narrowed (IQR: DAS28 1.8-9.3; MSUS 1.6-10.0), though there remained notable outliers in both groups (maximum 18 month CRP: DAS28 126, MSUS 67.0).



Graph 31 – Median CRP at each assessment time point

### 6.6.5 Change in Global Health 10cm VAS

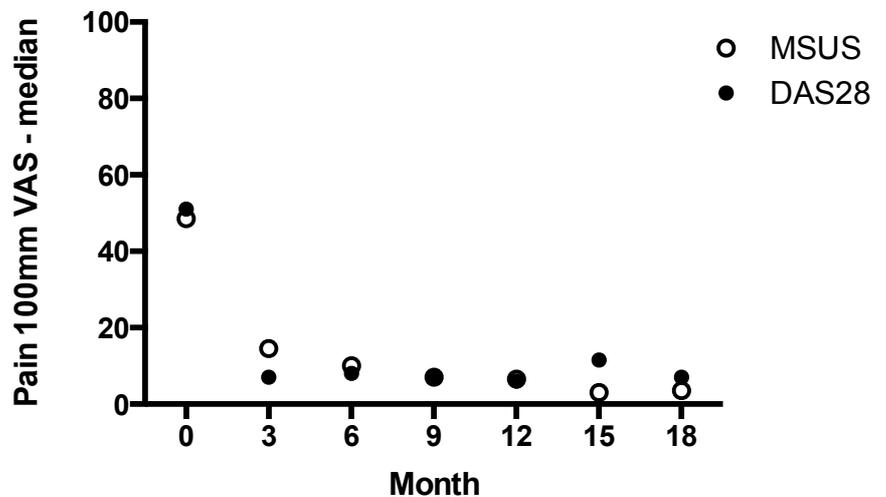
The median Global VAS at baseline was 54mm for the DAS28 group and 54.5mm for the MSUS group. After 18 months follow-up, the median Global VAS had fallen to 6mm in the DAS28 group and 2.0mm in the MSUS group (Graph 32). There were no statistically significant between group differences measured in Global VAS values at any of the assessment time points ( $p = 0.29-0.89$ ) or in the mean change from baseline ( $p = 0.65$ )



Graph 32 – Median Global Health 100mm Visual Analogue Score at each assessment time point

### 6.6.6 Change in Pain 100mm VAS

Participants in both groups reported similar levels of moderate to high pain 100mm visual VAS scores at baseline (median baseline pain 100mm VAS: DAS28 51 MSUS 48.5;  $p 0.30$ ) that had virtually normalised after 18 months (DAS28 7.0, MSUS 3.5,  $p 0.33$ ) (Graph 33). There were no significant between group differences in pain 100mm VAS at any time point ( $p = 0.23 - 0.90$ ), nor in the change from baseline ( $p = 0.68$ )



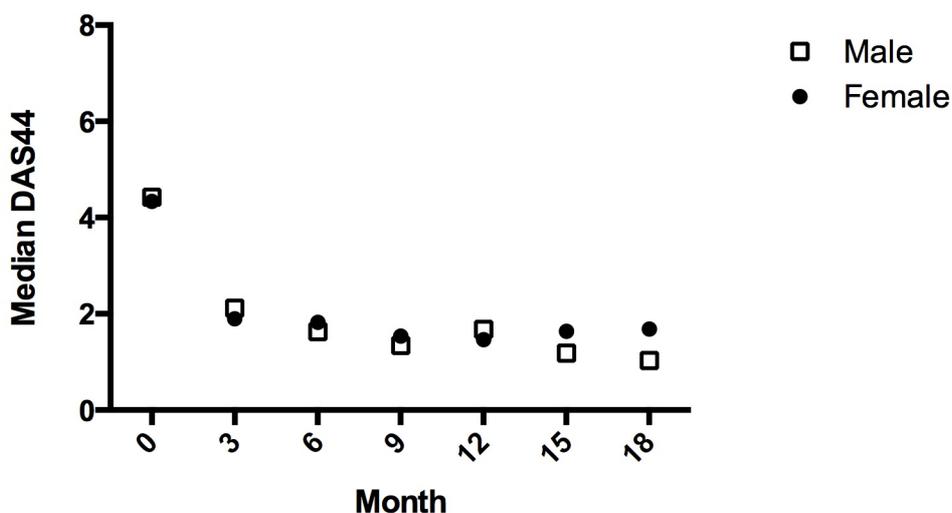
Graph 33 – Median Pain VAS at each assessment time point

## 6.7 Impact of gender on treatment response and remission rates

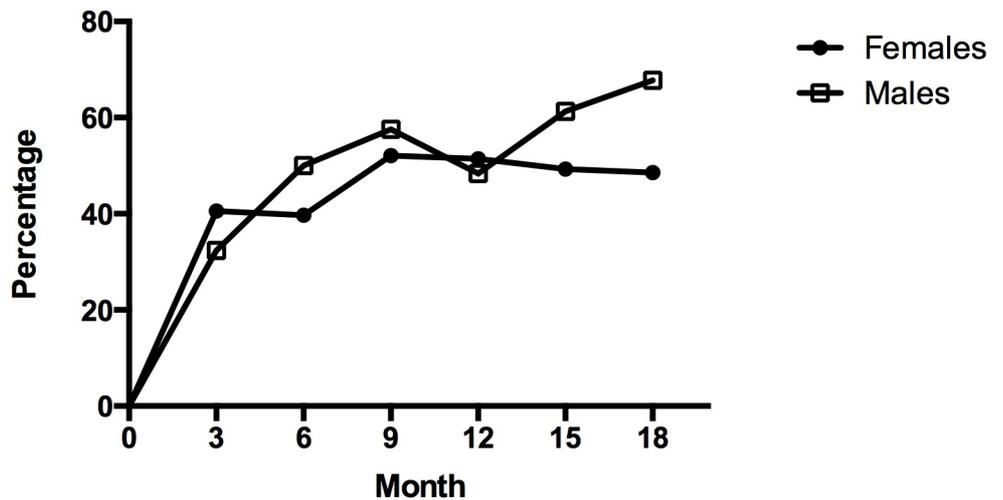
Female gender may have adverse prognostic implications, having been previously associated with lesser treatment responses, worse functional outcomes and greater rates of functional decline (38,40,45,320). The randomisation process allocated a significantly higher number of females into the DAS28 group (78% vs 60%,  $p$  0.031) and clearly this may have adversely impacted the measured outcomes of the DAS28 group and attenuated the true between group difference. To estimate the potential impact of gender on observed outcomes, gender-based analyses comparing the median DAS44, median HAQ and median change in DAS44 from baseline were conducted at a whole cohort and assessment group level.

### 6.7.1 Impact of gender on DAS44 and HAQ – whole cohort

Outcome data were pooled for the whole cohort and divided by gender rather than allocation group. At baseline, participants of both sexes reported similar scores for DAS44 and HAQ. Furthermore, both sexes exhibited significant improvements in DAS44 and HAQ over the whole of the follow-up period. At every time point, there was virtually no difference in the median DAS44 level (Graph 34). The pattern of mean change from baseline of DAS44 was similar for both sexes. There appeared to be a slightly higher rate of change measured in the Male group; however, there was no significant between group difference in the mean change from baseline measured at any time point ( $p$  0.21-0.97). Furthermore, linear regression analysis did not identify any significant differences in the equation of the line representing mean change in DAS44 from baseline (Slope of line: Females -0.031 vs Males -0.042,  $p$  0.46). DAS44 remission rates for both sexes were very similar for follow-up months 3 to 12 (Graph 35). For follow-up months 15 and 18 the proportion of participants attaining DAS44 remission tended to be higher in the male group (15 months Female 49% vs Male 61%; 18 months 49% vs 68%); however, neither of these differences were statistically significant ( $p$  0.29 and 0.09 respectively)

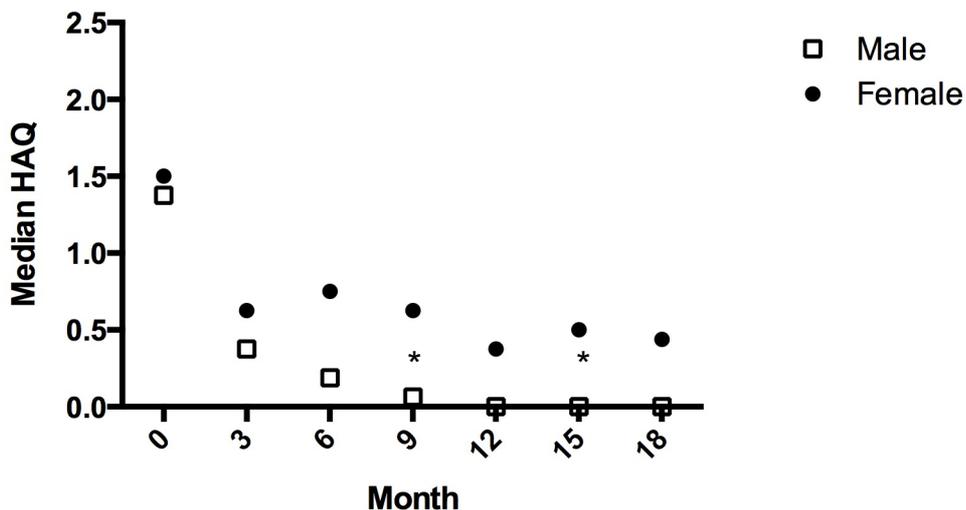


Graph 34 – Median DAS44 over follow-up period – segregated by gender



**Graph 35** - Percentage of participants attaining DAS44 remission, stratified by gender

Dividing participants by gender did identify a different pattern of improvement in the median HAQ over the follow-up period. Baseline HAQ levels were similar for both gender groups (median HAQ: Females 1.5 vs Males 1.375,  $p = 0.25$ ). However, at every other time point, male participants reported HAQ scores that were consistently numerically lower than their female counterparts (Graph 36). The between group difference in HAQ was statistically significant at months 9 (Females 0.63 vs Males 0.063,  $p = 0.02$ ) and months 15 (Females 0.50 vs Males 0.0,  $p = 0.038$ ). At every time point, the mean change from baseline of HAQ was consistently higher for male participants, though there was no difference measured in the mean overall change after 18 months follow-up (Females -0.88 vs Males -1.0,  $p = 0.66$ )

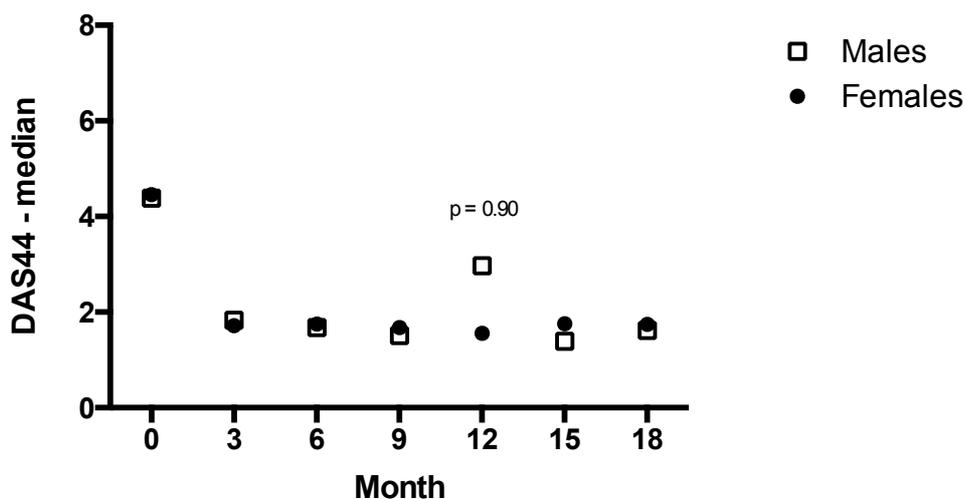


**Graph 36** – Median HAQ at each assessment time point, segregated by gender  
\*  $p < 0.05$

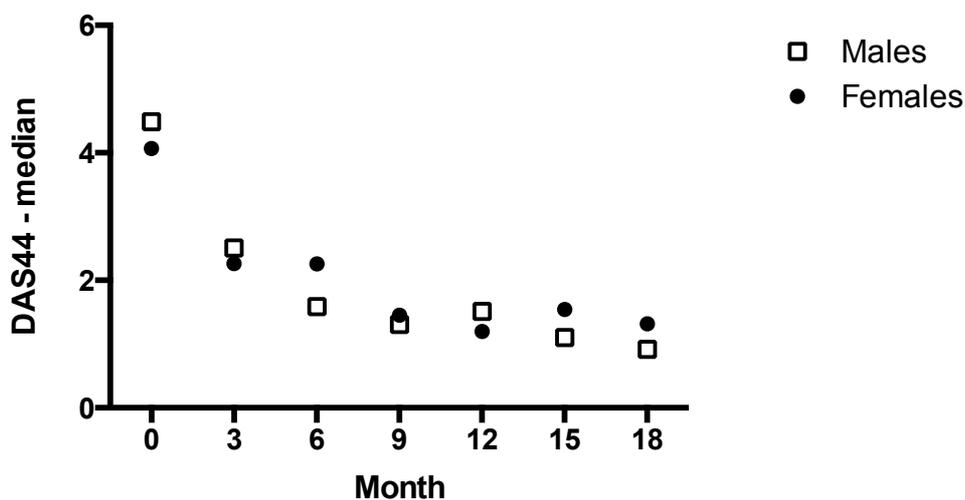
### 6.7.2 Impact of gender on DAS44 and HAQ – by assessment group

Subdividing randomisation groups by gender created small and numerically imbalanced comparator groups. In the DAS28 group there were 43 females and 14 males, whilst in the MSUS group there were 33 females and 21 males. Neither randomisation group exhibited significant

between-gender differences in the value of DAS44 at any of the assessment time points (Graph 37 – DAS28 group; Graph 38 – MSUS group) (p 0.13 – 0.98).

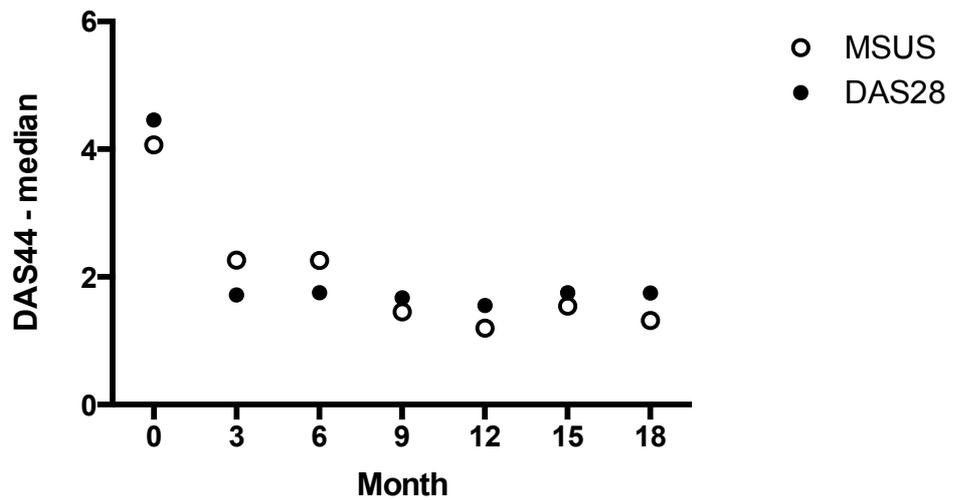


**Group 37 – DAS28 group:** median DAS44 at each assessment time point, segregated by gender (Males n = 14, Females n = 43)

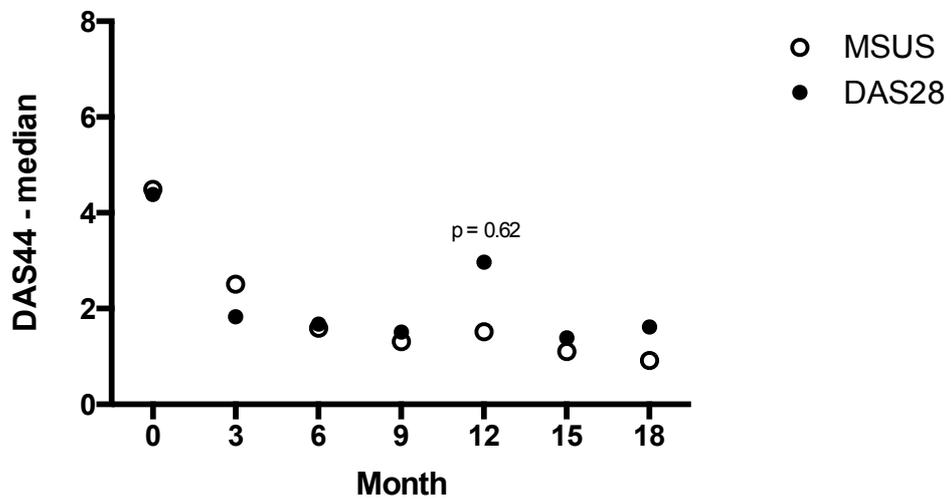


**Graph 38 – MSUS group –** median DAS44 at each assessment time point, segregated by gender (Males n = 21, Females n = 33)

Comparing the change in DAS44 between same-gender subgroups of the randomisation groups did not identify any significant differences in the pattern of change of DAS44 over the follow-up period. Female participants in the MSUS group exhibited slightly higher DAS44 scores at follow-up months 3 and 6 (median DAS44: DAS28 1.72 vs MSUS 2.26 and DAS28 1.75 vs 2.26 respectively) (Graph 39); however, neither of these differences reached statistical significance (p 0.72 and 0.61 respectively). There were no other notable differences over the rest of the follow-up period. Male participants in both groups exhibited virtually identical DAS44 scores for most of the follow-up period except month 12 when the DAS44 was numerically lower in the MSUS group (median DAS44: DAS28 2.97 vs MSUS 0.1.51) (Graph 40), however, this was not statistically significant (p 0.62)

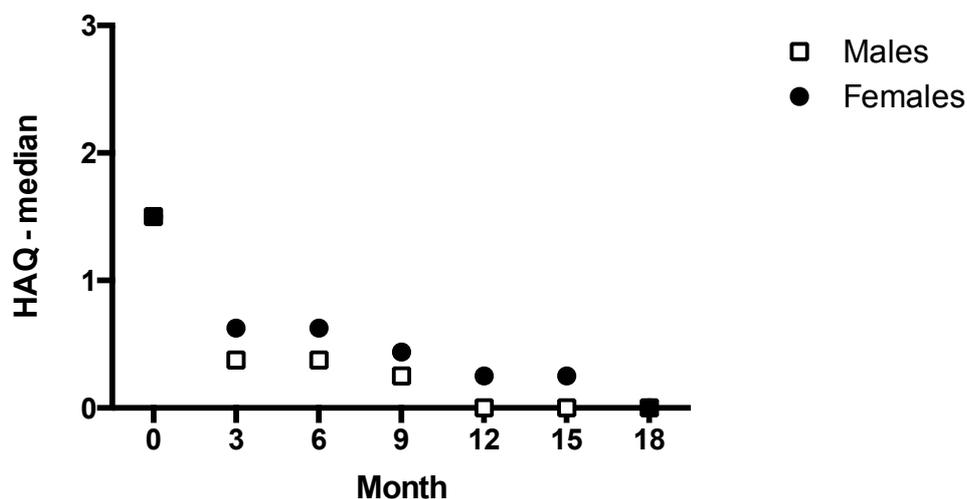


**Graph 39** – Median DAS44 for female participants, stratified by assessment group

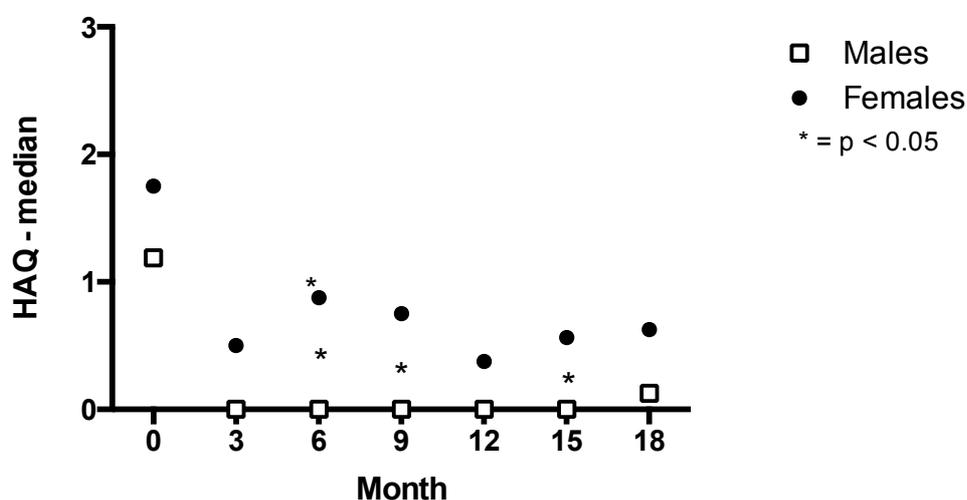


**Graph 40** – Median DAS44 for male participants, stratified by assessment group

In the MSUS group, there was no significant difference in the HAQ score reported by either sex at any of the time points (Graph 41) ( $p$  0.25-0.94). In the DAS28 group, where the gender distribution was less well balanced, the male participants consistently reported lower HAQ scores than the female participants with a statistically significant between sex difference evident at follow-up month 6 (median HAQ: Females 0.86 vs Males 0.0,  $p$  0.035), month 9 (0.75 vs 0.0,  $p$  0.032) and month 15 (median HAQ: Females 0.56 vs Males 0.0,  $p$  0.029) but not at any other time point (Graph 42)



**Graph 41 – MSUS group:** Median HAQ score at each assessment time point, segregated by gender



**Graph 42 – DAS28 group:** Median HAQ score at each assessment time point, segregated by gender  
\* = p < 0.05

## 6.8 Multiple Regression Analyses

To determine whether individual baseline variables were significant predictors of 18 month outcomes, multiple backward linear regression analyses were conducted using SPSS v19. The dependent variables were change in DAS44 between baseline and 18 months and 18 month HAQ. The independent (predictor) variables included: ordinal variables – randomisation group, gender, anti-CCP antibody status, smoking status; continuous variables – age, symptom duration, baseline HAQ, baseline DAS44, baseline CRP

*Change DAS44 baseline-18 months*

$R^2 = 0.33$ ,  $F = 17.2$ ,  $p < 0.001$

<b>Predictor Variable</b>	<b>Beta Co-efficient</b>	<b>p</b>
Baseline DAS44	-0.57	<0.001
Baseline CRP	-0.20	0.021

(Randomisation group, gender, anti-CCP antibody status, smoking status, age, symptom duration and baseline HAQ were not significant predictors in this model)

*18 month HAQ*

$R^2 = 0.24$ ,  $F = 34.8$ ,  $p < 0.001$

<b>Predictor Variable</b>	<b>Beta Co-efficient</b>	<b>p</b>
Baseline HAQ	0.49	<0.001

(Randomisation group, gender, anti-CCP antibody status, smoking status, age, symptom duration, baseline DAS44 and baseline CRP were not significant predictors in this model)

## **6.9 Discussion**

Based upon the available results it is possible to make the following conclusions about the potential role for MSUS disease activity assessment in the management of early RA:

1. In early RA, the proposed indications would lead to MSUS disease activity assessments being conducted during approximately 55% of consultations during months 3 to 18 of follow-up
2. The majority of MSUS disease activity assessments are indicated by DAS28 LDAS (89% of all MSUS assessments, 49% of all consultations). A small minority (11%) of MSUS disease activity assessments are indicated by DAS28 moderate disease activity but minimal synovitis
3. RA patients require a highly variable number of MSUS disease activity assessments when being managed using a MSUS guided step-up DMARD escalation regimen
4. MSUS assessments of synovitis activity (i.e. PD signal findings) seem to improve when RA patients are treated using a step-up DMARD escalation regimen
5. Using the proposed MSUS definitions of active disease, and a limited MSUS joint set, the majority (approximately 71%) of MSUS disease activity assessments agreed with the findings of the DAS28 . . . .
6. . . . however, in a substantial minority of occasions (approximately 29%), MSUS assessment provided disease activity findings that changed the perception of disease activity and supported an alternative treatment decision.

7. In DAS28 LDAS and remission, MSUS assessment provided additional disease activity information that supported further DMARD escalation in approximately 25% of examinations
8. In DAS28 moderate disease activity but minimal clinical synovitis, MSUS assessment identified active disease, and therefore supported DMARD escalation, in approximately 33% of examinations
9. From the proposed joint set, the radiocarpal and index MCP joints were most likely to exhibit PD signal
10. Early RA patients treated using either a DAS28 or MSUS-guided step-up DMARD escalation experience significant, and similar, improvements in measures of disease activity, functional ability, pain and the inflammatory response up until 18 months of follow-up.
11. MSUS guided DMARD escalation was not associated with any significantly better, or faster, improvements in ACR core set variables up until 18 months of follow-up . . . .
12. . . . but was associated with accrual of EULAR good responses and a higher rate of DAS44 remission after 18 months
13. MSUS guided DMARD escalation may be associated with a greater improvement in functional ability (HAQ)
14. Both DAS44 and ACR/EULAR definitions of remission demonstrated similar patterns of change in remission over the follow-up period. The between group differences were less marked when the more stringent ACR/EULAR Boolean definition was applied compared to the DAS44 threshold definition

Thus, these results seem to suggest that MSUS disease activity assessment might achieve additional benefits in some patients. Specifically, by increasing the likelihood of attaining remission, MSUS-guided DMARD escalation is in line with recent guidelines advocating that treatment targets should be remission (111,124,129,350). However, whilst attaining a marginally lower level of disease activity (i.e.  $DAS44 < 1.6$  vs  $DAS44 < 2.6$ ) could theoretically be associated with better functional and radiographic outcomes this has yet to be demonstrated in a clinical trial. The results also suggest that the MSUS group experienced a slightly greater overall improvement in functional ability. Indeed, by completion the majority of the MSUS group reported HAQ scores that were essentially normal and, arguably, improving functional ability should be a very important goal of treatment. Whilst there was no statistically significant between group difference in the mean change in HAQ the difference (0.23) was greater than the minimally clinically important difference (0.22) so it may be that the study was underpowered to detect a significant difference in change in HAQ.

Virtually identical clinical outcomes were achieved by the DAS28 group without having to use quite such aggressive DMARD combinations nor conduct time consuming MSUS examinations. If MSUS were to be incorporated into routine clinical practice, clinicians would need to consider whether the additional benefits experienced by patients were sufficient to justify the additional time (and training) required to conduct the assessment. Unfortunately, it is not yet possible to fully answer this important question, since it has not yet been shown that regular MSUS assessment is

associated with significantly better radiographic outcomes, nor whether the small increase in remission rates translate into significantly better medium-long term benefits. If MSUS driven DMARD escalation is not shown to significantly improve medium-long term outcomes, it is quite possible that its impact on clinic appointments may outweigh its relatively modest benefits. Particularly when very similar benefits can still be gained through DAS28 driven DMARD escalation. In this context, MSUS assessment of global disease activity may remain a useful additional method of quantifying global disease activity, to be considered for patients in whom there may be doubt about the accuracy of clinical disease activity assessment

Over the course of the follow-up period, patients in both assessment groups experienced a significant improvement in all of the disease activity measures. For each ACR core set variable, the greatest rate of improvement was observed between baseline and follow-up month 3. That is, the intervention that had the greatest impact on disease activity was the commencement of first DMARD and the initial, early administration of corticosteroids. In both groups, the values for each ACR core set variable remained relatively static between follow-up months 3 and 18. Routinely incorporating MSUS assessment into global disease activity assessment did not result in the MSUS assessment group exhibiting significantly better (i.e. lower) measures of either disease activity, functional ability or the acute phase response at any point during the follow-up period. In fact, the summary values of each ACR core set variable were similar, if not virtually identical, at each assessment time point and both groups exhibited very similar patterns of improvement over the whole of the follow-up period. Whilst it is not (yet) possible to specifically describe either the DMARD or corticosteroid treatment exposure of each group over the course of the study it is highly likely that the MSUS group will have received more aggressive therapy than the DAS28 group. It is likely that this would have been evident as both an increased, and earlier, use of combination DMARD therapy and an increased use of etanercept therapy either at an earlier point in the disease course or at a level of disease activity not normally associated with requiring biologic therapy. However, this greater intensity of DMARD therapy over the duration of the follow-up period did not seem to translate into significantly better values of ACR core set variables at any point. The condition of individual patients may have improved significantly; however, the additional time spent conducting MSUS disease activity assessments in addition to the DAS28 may not necessarily lead to better clinical outcomes that can be measured at a group level

Even though the MSUS group did not exhibit significantly better measures of disease activity at any time point, there was some evidence that the group overall did accrue additional moderate benefits that were not evident in the DAS28 group. For example, for the MSUS group there was an increase in the proportion of participants attaining EULAR good responses between months 6 and 18 (55% to 74%) that was statistically significant ( $p = 0.04$ ). By contrast, in the DAS28 group, there was no additional improvement in the rate of EULAR good responses between months 6 and 18 and, if anything, there was actually a slight decline (month 6 = 67%, month 18 = 63%).

DAS44 remission rates were similar for both groups between follow-up months 3 to 15; however, by month 18, the MSUS group were exhibiting a statistically higher rate of DAS44 remission than

the DAS28 group (66% vs 43%,  $p$  0.028). The proportion of DAS28 group participants who attained DAS44 remission remained static overall between follow-up months 3 and 18 (42% to 43%), and if anything declined between months 6 and 18 (48% to 43%). By contrast, the MSUS group demonstrated an overall increase in the proportion of participants achieving DAS44 remission between months 3 and 18 (34% to 66%,  $p$  0.0016). From the shape of the associated graph (Graph 24), it is tempting to suggest that this was a step-wise incremental increase. However, there was no statistical difference in the DAS44 remission rates between DAS28 and MSUS groups at any assessment point until month 18. Excepting month 6, the DAS44 remission rates for both groups are closely matched and follow similar patterns. It is only really at month 18 that there is any divergence evident between the groups. Participants within the MSUS group were most likely to commence etanercept within the last few months of the follow-up period and this may partly explain some of the divergence in month 18 DAS44 remission rates. However, a longer follow-up would be needed to determine whether there is continued divergence in the remission rates and whether separation in any of the other ACR core sets becomes apparent. Taken altogether, it is possible that MSUS assessment assisted a subset of MSUS group participants to achieve a tighter level of disease control than may have been possible had therapy been guided by DAS28 alone. This is best suggested, by the continued increase in DAS44 remission rates, accumulation of participants with a EULAR good response and the MSUS group experiencing a slightly greater overall mean change in HAQ. Whilst the randomisation process lead to groups with unequal gender distributions, it is unlikely that this could have significantly biased the observed results since: 1. within each randomisation group, male and female participants exhibited similar patterns of change in DAS44, 2. there were no significant differences in the pattern of change of DAS44 between female participants of each group and 3. there were no significant differences in HAQ values at any time point between male and female subgroups.

Previous studies have shown that the persistence of PD signal in asymptomatic patients predicts the risk of future disease flare (217). The MSUS group treatment strategy proposed by this research used the persistence of PD signal as a marker of continuing disease activity and the primary trigger for DMARD escalation. Whilst this did not lead to the MSUS group achieving significantly higher overall DAS44 remission rates (except at month 18), it may have lead to an overall lower rate of PD signal (i.e. more complete suppression of synovitis), a lower likelihood of fluctuations in disease activity and an increased likelihood of MSUS group participants achieving DAS44 remission on a higher number of occasions. Long term follow-up is required of both assessment groups to determine whether the subtle short term difference translate into an altered medium-long term disease course for either group (i.e. lesser rate of acute flare between 18 and 60 months). However, this partly assumes that patients receiving MSUS guided DMARD therapy will also exhibit lower long term levels of PD signal than those receiving DAS28 guided therapy. This assumption cannot be supported since there is no comparator PD signal data available from the DAS28 group. Previous observational studies have suggested that PD findings do improve in response to escalations of DMARD and biologic therapy (351,352). The available results did show improvements in both the total PD score and PD joint count; however, it is not possible to determine whether this is related to the additional influence of MSUS on treatment intensity or

whether similar changes would have been evident in the DAS28 group. Ideally, future studies of the impact of MSUS guided DMARD therapy would also collect MSUS findings data for their control arms too. Even if these data were not used to influence treatment decisions it would serve as a useful comparator to determine whether MSUS guided DMARD therapy results in lesser rates of PD signal positivity and to determine how this relates to corresponding changes in disease activity

So far, the reported results have taken a short-term view of the participant's response to step-up DMARD therapy. It is possible that the follow-up period was too short, and the treatment response in the DAS28 group too good, to expect a between group difference to become evident within the time available. Further, an important predictor of long term prognosis, and therefore an important marker of the likely long term efficacy of the treatment strategy, is the development of radiographic erosions (3). Even though there was no significant difference in clinical response observed it is possible that the more aggressive treatment regimen received by the MSUS assessment group was associated with a lesser rate of progression in radiographic damage. This relationship will be clearer once the plain x-ray outcomes have been formally graded. Equally, comparisons of the change in MRI RAMRIS outcomes between baseline and 18 months will provide an additional measure of the impact of each group's DMARD therapy on several very sensitive, and often sub-clinical, measures of RA activity. It is possible that grading of the MRI RAMRIS images will demonstrate favourable changes in the MSUS group with a lesser change in the RAMRIS erosion score (i.e. lesser destructive progression), a greater change in the RAMRIS synovitis score (i.e. a greater treatment response), and a greater change in the RAMRIS bone marrow oedema score (i.e. a lesser risk of developing future erosions). In this way, it may be possible to demonstrate that MSUS-guided step-up DMARD therapy is associated with favourable short-medium term radiological outcomes that are likely to be associated with a lesser long-term risk of developing joint destruction and chronic disability. Thus, even though there appears to be no short-medium term benefit in clinical response it may yet be possible to demonstrate a potential benefit in medium-long term radiological benefits that justifies the routine use of MSUS monitoring of global disease activity in early RA.

## **7. The Relationship Between RA Phenotype and Gene Expression Profile**

To increase cohort size and microarray sensitivity expression profile datasets for DAS28 and MSUS assessment groups were pooled. Certain participants were excluded if they had either withdrawn from on going follow-up or had not provided sufficient information for analysis at a particular time point. Comparison groups were formed by segregating participants into groups based upon clinical, or treatment response phenotype as previously described in sections 2.7.2 and 3.8.2. However, as the analysis progressed additional comparisons were conducted that were either based upon the outputs of pre-planned comparisons and/or were an attempt to maximise the likelihood of identifying a significant finding.

When comparing expression levels of individual genes between different clinical groupings an adjusted p value (false discovery rate) less than 5% was considered significant. In many comparisons the lowest adjusted p value observed was actually greater than 5%. In these instances, the acceptable adjusted p value for that comparison was adjusted arbitrarily depending upon the values returned for that comparison. However, there were some comparisons where the adjusted p value were so high (greater than 20%) it was unlikely that it would be possible to confidently describe a significant between group difference in gene expression. In the following sections, comparisons that identified significant between group differences in gene expression will be presented in conjunction with detailed gene lists taken directly from the Limma comparison. Conversely, for the sake of space and brevity, comparisons that failed to identify significant evidence of differential gene expression will be presented with a description of the best observed (i.e. lowest) adjusted p value but not the full gene lists.

## **7.1 Clinical Features of Gene Expression Analysis Cohort**

79 participants provided PAXgene RNA samples and clinical data for the gene expression analysis. 76 of these participants (96%) fulfilled the 2010 ACR classification criteria for RA. For 73 participants (92%) methotrexate was the first choice of DMARD monotherapy, the remaining 6 patients all received sulfasalazine monotherapy. Forty participants (51%) were randomised to the DAS28 assessment group and 39 (49%) were randomised to the MSUS assessment group. Table 41 summarises the baseline features of the cohort. Overall, the baseline features of the gene expression analysis cohort closely matched the baseline features of the whole TaSER cohort (also shown in Table 41).

Overall, the transcriptomic analysis cohort experienced a good response to step-up DMARD escalation therapy (Table 42). There was a clear numerical fall in DAS28 between baseline and month 3 and a lesser fall between month 3 and month 18. There were a statistically significant falls in DAS28 between baseline and month 3 (paired t test:  $p < 0.0001$ ), baseline and month 18

(paired t test:  $p < 0.0001$ ) and between month 3 and month 18 (paired t test:  $p < 0.0001$ ).

Furthermore, the mean improvement in DAS28 between baseline and month 3 and baseline and month 18 (mean change DAS28: -2.1 and -2.7 respectively) were both greater than the minimum clinically important change (1.2). Statistically significant improvements in ESR and CRP were also measured between baseline and month 3 and baseline and month 18 (paired t test:  $p < 0.0001$  for each comparison).

	<b>Transcriptomic Analysis Cohort (n=79)</b>	<b>Whole TaSER Cohort (n=111)</b>
<b>Female Sex – n (%)</b>	54 (68%)	68%
<b>Age (years)</b>	56 (SD 13)	56 (SD 13)
<b>Disease Duration (months)</b>	5.3 (SD 3.1)	5.3 (SD 3.0)
<b>Rheumatoid Factor Positive - n (%)</b>	51 (65%)	67%
<b>Anti-CCP Antibody Positive - n (%)</b>	43 (54%)	60%
<b>DAS28</b>	5.0 (SD 1.1)	5.0 (SD 1.1)
<b>HAQ</b>	1.5 (SD 0.80)	1.5 (SD 0.80)
<b>ESR</b>	36 (SD 26)	39 (SD 35)
<b>CRP</b>	42 (SD 55)	37 (SD 49)
<b>Plain X-ray erosions – n (%)</b>	26 (33%)	29%

**Table 41** – Baseline features of transcriptomic analysis cohort in comparison to the whole research cohort. Unless otherwise stated values are means (standard deviation)

	<b>DAS28</b>	<b>Mean change from baseline DAS28</b>	<b>ESR</b>	<b>CRP</b>
<b>Baseline</b>	5.0 (SD 1.1)		36 (SD 26)	42 (SD 55)
<b>Month 3</b>	2.9 (SD 1.2)	-2.1 (SD 1.4)	21 (SD 21)	14 (SD 26)
<b>Month 18</b>	2.3 (SD 1.0)	-2.7 (SD 1.5)	15 (SD 12)	9 (SD 12)

**Table 42** – Disease activity measures at baseline, 3 months and 18 months for the transcriptomic analysis cohort. All values are mean (standard deviation)

Comparator groups were formed based upon the available DAS28 data, though this was recorded in a manner that was not blinded to the participant's randomisation group. Towards the end of the data analysis period, the DAS44 data also became available. Whilst the DAS44 data could not be used to form comparator groups, it could be used to demonstrate that the gene expression analysis cohort had experienced a similar level of disease activity and treatment response to the whole TaSER cohort (Table 43). The mean values of DAS44 for the gene expression and TaSER cohorts were similar at each of the time points and the mean change from baseline for both cohorts were similar between baseline and month 3 and baseline and month 18.

The proportion of gene expression analysis participants categorised as different DAS44 disease activity states are shown in Table 44

	DAS44		Mean change from baseline of DAS44	
	Transcriptomics Cohort	Whole Cohort	Transcriptomics Cohort	Whole Cohort
<b>Baseline</b>	4.5 (SD 1.2)	4.4 (SD 1.1)		
<b>Month 3</b>	2.4 (SD 1.3)	2.2 (SD 1.3)	-2.1 (SD 1.4)	-2.1 (SD 1.3)
<b>Month 18</b>	1.7 (SD 1.2)	1.7 (SD 1.1)	-2.7 (SD 1.5)	-2.6 (SD 1.5)

**Table 43** – DAS44, and change from baseline of DAS44, at baseline, 3 months and 18 months for transcriptomics analysis cohort and whole research cohort. Values are mean (standard deviation)

	Baseline	Month 3	Month 12	Month 18
<b>High</b> DAS44 >3.7	57 (72%)	13 (17%)	7 (10%)	5 (7%)
<b>Moderate</b> 2.4 ≤ DAS44 <3.7	22 (28%)	23 (30%)	15 (21%)	17 (24%)
<b>Low</b> DAS44 <2.4	0	41 (53%)	50 (69%)	49 (69%)
<b>Remission</b> DAS44 <1.6	0	28 (26%)	39 (54%)	39 (55%)

**Table 44** – Number (percentage) of transcriptomic analysis cohort attaining DAS44-defined disease activity states at each time point

## 7.2 Baseline Phenotypic Groupings

Participants were divided into phenotypic comparator groups based upon their presenting characteristics. Between group comparisons of differences in gene expression were then conducted using the baseline, pre-treatment gene expression profiles. Pre-treatment expression profiles were chosen to minimise any influence DMARD therapy might have upon the observed gene expression patterns

### 7.2.1 Gender

Segregating participants by sex identified the highest number of differentially expressed genes. In total, 66 genes exhibited differential expression between female and male groups (FDR range 2.26E-34 – 0.0458). In females, 25 genes were down-regulated and 36 genes were up-regulated relative to males. Genes were associated with a range of different loci (chromosomes 1, 2, 3, 5, 6, 10, 11, 12, 14, 18, 19 and 20). All 14 genes associated with the X chromosome were up regulated whereas all 16 genes associated with the Y chromosome were down regulated. Four genes (EIF1AY, RPS4X, UTY, EIF1AX) were each identified at more than one microarray location. Additionally four probe identifiers (LOC100133662, 38961, LOC644670, LOC647322) were not associated with recognised genes by the online Genes database

([www.ncbi.nlm.nih.gov/gene](http://www.ncbi.nlm.nih.gov/gene)). Table 45 (see over) lists the expression levels and chromosomal

location of all genes that exhibited differential expression between females and males in this analysis

Illumina Gene Symbol	Gene Name	log Fold Change	Adjusted p Value	Chromosomal Location
RPS4Y1	ribosomal protein S4, Y-linked 1	-3.90	2.3E-34	Yp11.3
LOC100133662 <sup>α</sup>	?	-2.56	2.1E-30	X
EIF1AY* <sup>1</sup>	eukaryotic translation initiation factor 1A, Y-linked	-1.65	2.6E-27	Yq11.223
JARID1D	lysine (K)-specific demethylase 5D	-1.20	2.6E-27	Yq11
PRKY	Protein kinase, Y-linked	-1.20	7.0E-27	Yp11.2
CYORF15A	chromosome Y open reading frame 15A	-0.67	3.0E-23	X
XIST	X inactive specific transcript	1.26	1.5E-22	Xq13.2
EIF1AY* <sup>1</sup>	eukaryotic translation initiation factor 1A, Y-linked	-0.70	2.7E-21	Yq11.223
RPS4Y2	ribosomal protein S4, Y-linked 2	-1.13	1.3E-19	Yq11.223
CYORF15B	taxilin gamma 2, pseudogene	-0.38	1.0E-15	Yq11.222
PRKX	Protein kinase, X-linked	0.35	2.4E-12	Xp22.3
UTX	lysine (K)-specific demethylase 6A	0.37	1.0E-11	Xp11.2
LOC391777	ribosomal protein S4X pseudogene 6	0.38	6.6E-09	5p13.2
EIF2S3	eukaryotic translation initiation factor 2, subunit 3 gamma	0.37	3.3E-08	Xp22.2-p22.1
ZFY	Zinc finger protein, Y-linked	-0.23	1.4E-07	Yp11.3
RPS4X * <sup>2</sup>	ribosomal protein S4, X-linked	0.54	1.9E-06	Xq13.1
UTY * <sup>3</sup>	ubiquitously transcribed tetratricopeptide repeat containing, Y-linked	-0.23	4.0E-06	Yq11
RPS4X * <sup>2</sup>	ribosomal protein S4, X-linked	0.53	4.3E-06	Xq13.1
UTY * <sup>3</sup>	ubiquitously transcribed tetratricopeptide repeat containing, Y-linked	-0.24	5.4E-06	Yq11
ZBED1	zinc finger, BED-type containing 1	-0.27	1.3E-05	Xp22.33
38961 <sup>α</sup>	?	0.42	0.00018	
CA5B	carbonic anhydrase VB, mitochondrial	0.33	0.00077	Xp21.1
TMSB4Y	thymosin beta 4, Y-linked	-0.18	0.00092	Yq11.221
SOCS2	suppressor of cytokine signaling 2	0.34	0.00096	12q
LOC441550	<b>ribosomal protein S4X pseudogene 11</b>	0.27	0.0023	10p11.22
PURA	purine-rich element binding protein A	0.27	0.0026	5q31
EIF1AX * <sup>4</sup>	eukaryotic translation initiation factor 1A, X-linked	0.21	0.0035	Xp22.12
EIF1AX * <sup>4</sup>	eukaryotic translation initiation factor 1A, X-linked	0.28	0.0036	Xp22.12
ZRSR2	zinc finger (CCCH type), RNA-binding motif and serine/arginine rich 2	0.22	0.0038	Xp22.1
ZNF548	Zinc finger protein 548	0.23	0.0044	19q13.43
CSF2RA	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	-0.23	0.0068	Xp22.32 and Yp11.3
HDHD1A	haloacid dehalogenase-like hydrolase domain containing 1A	0.22	0.0070	18
LOC220433	ribosomal protein S4X pseudogene 16	0.34	0.0072	13q14.3
LOC644670 <sup>α</sup>	?	0.23	0.0088	
LOC390183	ribosomal protein S4X pseudogene 13	0.28	0.0088	11q12.1
DOCK10	dedicator of cytokinesis 10	0.31	0.0090	2q36.2

**Table 45** - Genes exhibiting differential expression in baseline samples between female and male RA patients. Ranked by adjusted p value  
Adjusted p values < 0.05 are considered significant  
\*denotes duplicate gene; α denotes unrecognised gene

Illumina Gene Symbol	Gene Name	log Fold Change	Adjusted p Value	Chromosomal Location
CXORF45	UDP-N-acetylglucosaminyltransferase subunit	0.24	0.0090	Xq23
TFRC	transferrin receptor (p90, CD71)	0.23	0.0096	3q29
SLITRK3	SLIT and NTRK-like family, member 3	-0.13	0.011	3q26.1
LOC554203	JPX transcript, XIST activator (non-protein coding)	0.22	0.011	Xq13.2
USP9Y	ubiquitin specific peptidase 9, Y-linked	-0.15	0.012	Yq11.2
BCORL2	BCL6 corepressor pseudogene 1	-0.14	0.013	Yq11.222
ZNF248	zinc finger protein 248	0.18	0.014	10p11.2
LOC647322 <sup>α</sup>	?	0.20	0.016	
LRRC58	leucine rich repeat containing 58	0.21	0.02	3q13.33
TNNI2	troponin I type 2 (skeletal, fast)	-0.17	0.022	11q15.5
CLTB	clathrin, light chain B	-0.16	0.023	5q35
BACH2	BTB and CNC homology 1, basic leucine zipper transcription factor 2	0.28	0.023	6q15
SLC20A1	solute carrier family 20 (phosphate transporter), member 1	0.19	0.025	2q13
ZNF512	zinc finger protein 512	0.20	0.025	2p23
IL7R	Interleukin 7 receptor	0.53	0.026	5p13
BCL2	B-cell CLL/lymphoma 2	0.35	0.030	18q21.3
EDG1	sphingosine-1-phosphate receptor 1	0.31	0.030	1p21
LRIG1	leucine-rich repeats and immunoglobulin-like domains 1	0.18	0.031	3p14
ITK	IL2-inducible T-cell kinase	0.34	0.031	5q31-q32
NANP	N-acetylneuraminic acid phosphatase	-0.14	0.040	20p11.1
C6ORF190	thymocyte selection associated	0.35	0.042	6q22.33
ZCCHC24	zinc finger, CCHC domain containing 24	-0.15	0.042	10q22.3
DRD3	Dopamine receptor D3	-0.18	0.045	3q13.3
ZNF540	Zinc finger protein 540	0.17	0.046	19q13.12
CD2	CD2 molecule	0.36	0.046	1p13.1
LBH	limb bud and heart development	0.32	0.046	2p23.1
MAPK14	mitogen-activated protein kinase 14	-0.26	0.046	6p21.3-p21.2
NAP1L1	nucleosome assembly protein 1-like 1	0.26	0.046	12q21.2
MBIP	MAP3K12 binding inhibitory protein 1	0.14	0.046	14q13.3
LOC387820	DnaJ (Hsp40) homolog, subfamily B, member 7 pseudogene	-0.13	0.046	11q24.3

**Table 45 (continued):** Genes exhibiting differential expression in baseline samples between female and male RA patients. Ranked by adjusted p value  
Adjusted p values < 0.05 are considered significant  
\*denotes duplicate gene; α denotes unrecognised gene

## 7.2.2 Rheumatoid Factor Status

Segregating patients based upon RF status identified 3 genes which exhibited differential expression between rheumatoid factor positive and negative groups when the adjusted p value threshold was less than 0.05. However, the microarray location (probe ID 7650093) with the lowest adjusted p value was not associated with a recognised gene. Increasing the adjusted p

value threshold to 0.15 identified a further 5 differentially expressed genes. Of these, one gene duplicated a previously identified gene (NAIP) and one gene (LOC648984) was not associated with a recognised chromosomal location. All five of the remaining genes were up-regulated in baseline samples of rheumatoid factor positive participants. Two up-regulated genes duplicated the same chromosomal location (NAIP and LOC728519, Chr 5q13.2) Table 46 lists the expression levels and chromosomal location of all genes that exhibited differential expression using the adjusted value threshold of less than 0.15

Illumina Gene Symbol	Gene Name	log Fold Change	Adjusted p Value	Chromosomal Location
NA (7650093)	Not recognised	-0.17	0.0092	
SRI	Sorcin	0.18	0.035	7q21.1
NAIP	NLR family, apoptosis inhibitory protein	0.30	0.04	5q13.2
LOC648984	Similar to Baculoviral IAP repeat-containing protein 1	0.38	0.13	Not known
LOC728519	NLR family, apoptosis inhibitory protein pseudogene	0.39	0.13	5q13.2
AP1S1	Adaptor-related protein complex 1, sigma 1 subunit	0.15	0.13	7q22.1
BCS1L	BC1 (ubiquinol-cytochrome c reductase) synthesis like	0.17	0.13	2q33

**Table 46:** Genes exhibiting differential expression between baseline samples of rheumatoid factor positive and negative patients  
Adjusted p values < 0.15 are considered significant

### 7.2.3 Anti-CCP Antibody Status

Segregating participants by baseline anti-CCP antibody status did not identify significant between group differences in gene expression. Whilst many genes exhibited highly significant p values during t-testing (down to  $p = 2.24E-05$ ), correcting for multiple testing returned consistently high adjusted p values (lowest adjusted p value = 0.51). Since expression of anti-CCP antibodies is strongly associated with tobacco smoke exposure, which in turn can also influence gene expression (353,354), the comparison of gene expression between anti-CCP comparator groups was rerun whilst also adjusting for the participant's smoking status. However, this corrected analysis also failed to identify any new differences in gene expression (lowest adjusted p value = 0.999).

### 7.2.4 Baseline Erosive Status

The presence or absence of erosions at first presentation was not associated with significant between group differences in gene expression. Simple t-testing returned highly significant p values for multiple genes; however, correcting for multiple testing suggested that there were no significant between group differences (lowest adjusted p value = 0.999)

### 7.2.5 Smoking Status

The impact of smoking status on gene expression in baseline samples was examined by dividing participants into groups based upon whether they were current, former or never smokers. Gene expression in each group was then compared in a step-wise fashion against gene expression in each of the other two smoking status groups. An adjusted p value threshold was initially set at 0.05. Current smokers exhibited up-regulation of the gene LRRN3 in relation to both never smokers (adjusted p value = 0.0021) and former smokers (adjusted p value = 0.0031). In both comparisons, there was evidence of LRRN3 up-regulation at 2 separate microarray locations. There was no evidence of differential LRRN3 expression between former and never smokers. Former smokers did exhibit up regulation of HIATL1 (adjusted p value = 0.040) in relation to never smokers. Furthermore, using an adjusted p value threshold of 0.15, former smokers also exhibited up-regulation of gene MBD2 in relation to never smokers. Table 47 describes the differences in gene expression levels between the different smoking status groups.

Illumina Gene Symbol	Gene Name	log Fold Change	Adjusted p Value	Chromosomal Location
<b>Current smoker vs Never Smoked</b>				
LRRN3	Leucine rich repeat neuronal 3	0.74	0.0021	7q31.1
LRRN3	Leucine rich repeat neuronal 3	0.43	0.0070	7q31.1
<b>Current smoker vs Former smoker</b>				
LRRN3	Leucine rich repeat neuronal 3	0.82	0.0031	7q31.1
LRRN3	Leucine rich repeat neuronal 3	0.48	0.0075	7q31.1
<b>Former smoker vs Never smoked</b>				
HIATL1	Hippocampus abundant transcript-like 1	0.25	0.040	9q22.32

**Table 47** - Genes exhibiting differential expression in baseline samples between current, former and never smokers  
Adjusted p values < 0.15 are considered significant  
\*denotes duplicate gene

## 7.3 Gene Expression Profiles as Measures of Global Disease Activity

Participants were divided into comparator groups based upon the numerical value of the DAS28 at baseline and three months. Between group comparisons of differences in gene expression were then conducted using the corresponding mRNA expression profiles. The presumption being that different disease activity states might associate with specific gene expression patterns. In order to try and maximise the sensitivity of the comparisons the numerical value of DAS28 was used in a variety of different permutations either as a stand alone value or in relation to predefined, disease activity classification thresholds.

### **7.3.1 Baseline Gene Expression Profile and Baseline Disease Activity – DAS28 Quartiles**

Categorising participants into groups based upon baseline disease activity level failed to identify significant differences in baseline gene expression profiles between those with the highest and lowest levels of disease activity. Participants were subdivided into groups based upon baseline DAS28 quartiles. To increase the likelihood of identifying clear between group differences, gene expression comparisons were limited to the upper and lower quartile groups. There were 19 participants in the lower quartile group where the DAS28 scores ranged from 2.83-3.95 (median 3.57). There were 18 participants in the upper quartile group where the DAS28 scores ranged from 5.85-8.21 (6.48). Mann-Whitney U test confirmed that there was a statistically significant difference in the distribution of the DAS28 scores between the upper and lower quartile groups ( $p < 0.0001$ ). However, despite the evident differences in clinical phenotype, there were no significant differences in gene expression profiles (lowest adjusted  $p$  value = 0.40)

### **7.3.2 Three Month Gene Expression Profile and Three Month Disease Activity – DAS28 Quartiles**

Categorising participants into groups based upon 3 month disease activity also failed to identify any differences in gene expression profile between the groups. Participants were divided into groups based upon quartiles of their 3 month DAS28 score and between group comparisons of gene expression profile were limited to the upper and lower quartile groups. There were 19 participants in the lower quartile group (DAS28 range: 1.049 – 2.025, median 1.73) and all fulfilled clinical remission criteria. There were 18 participants in the upper quartile group (DAS28 range: 4 – 6.3, median 4.46) with disease activity ranging from moderate to high. Mann Whitney U test confirmed that the distributions of DAS28 scores for each group were statistically distinct ( $p < 0.0001$ ). However, despite the evident differences in clinical phenotype, no significant differences in gene expression profile could be demonstrated between the groups (lowest adjusted  $p$  value = 0.78).

### **7.3.3 Three Month Gene Expression Profile and Three Month Disease Activity – DAS28 Thresholds**

Participants were divided into comparator groups based upon the numerical value of their 3 month DAS28 score in relation to existing DAS28-based definitions of clinical disease activity (e.g. high / moderate / low / remission). The 3 month expression profiles of the high disease activity and clinical remission groups were chosen for comparison since these were likely to be the most widely separated clinical phenotypes and therefore also most likely to exhibit differences in gene expression. However, the eventual groups were significantly imbalanced in size and therefore any attempted comparisons would have been significantly biased. Thirty five patients fulfilled clinical remission criteria (DAS28 range: 1.049 – 2.599, median 2.008) whereas only 4 patients displayed high disease activity (DAS28 range: 5.474 – 6.3).

### 7.3.4 Three Month Gene Expression Profile and Clinical Threshold for DMARD Escalation

To determine whether or not specific gene expression profiles were associated with an adequate treatment response after 3 months of DMARD monotherapy participants were subdivided into comparator groups depending on whether or not their 3 month DAS28 fell below the strategic threshold of 3.2. Fifty-three participants demonstrated a DAS28 less than 3.2 (range 1.049 – 3.161, median 2.35) and therefore would not ordinarily have been considered for DMARD escalation. Twenty five patients demonstrated a DAS28 score greater than 3.2 (i.e. range 3.205 – 6.3, median 3.70) and would have qualified for DMARD escalation. Mann-Whitney U testing confirmed that the distributions of DAS28 scores for both groups were statistically distinct ( $p < 0.0001$ ). Despite the differences in phenotype, there was only evidence of borderline differential expression of one single gene. The HS.515967 gene, which relates to a transcribed locus, was down regulated in patients who failed to achieve a 3 month DAS28 less than 3.2

Illumina Gene Symbol	Gene Name	log Fold Change	Adjusted p Value	Chromosomal Location
HS.515967	? – transcribed locus	-0.13	0.052	?

**Table 48** - Characteristics of a single gene which was down regulated in participants who failed to achieve 3 month DAS28 < 3.2

## 7.4 Baseline Gene Expression Profile as Predictor of Clinical Course

The relationship between baseline expression profiles and subsequent clinical outcomes was examined to determine whether baseline profiles had any predictive ability for short-medium term outcomes

### 7.4.1 Baseline Gene Expression Profile and Three Month Threshold for DMARD Escalation

To determine whether specific baseline gene expression patterns were present in those patients who were likely to require DMARD escalation (a surrogate for persistently active disease) after 3 months of DMARD monotherapy, baseline expression profiles were compared in participants grouped according to whether or not 3 month DAS28 fell below or above 3.2. Comparison groups were identical to those described in preceding section 7.3.4; however, on this occasion the baseline expression profiles were compared. No differentially expressed genes were associated with an adjusted p value threshold of less than 0.05. Using an adjusted p value threshold of 0.15, 3 genes were found to be up-regulated in baseline samples of participants exhibited a DAS28 greater than 3.2 after 3 months of DMARD monotherapy. The expression levels and chromosomal locations of these genes are described in Table 49.

Illumina Gene Symbol	Gene Name	log Fold Change	Adjusted p Value	Chromosomal Location
TBC1D22B	TBC1 domain family, member 22B	0.22	0.088	6p21.2
TMCC2	Transmembrane and coiled-coil domain family 2	0.27	0.088	1q32.1
USP46	Ubiquitin specific peptidase 46	0.14	0.088	4q12

**Table 49** - Genes which were up-regulated in baseline samples of patients with 3 month DAS28 greater than 3.2  
Adjusted p value threshold < 0.15

#### **7.4.2 Baseline Gene Expression Profile and Three Month Disease Activity Level – DAS28 Quartiles**

Baseline expression profiles were not associated with, and therefore are unlikely to be predictive of, participants subsequently achieving particular levels of disease activity. Participants were separated into quartile groups based upon their 3 month DAS28 score. Between group comparisons of differences in baseline expression profiles were then conducted. The composition of the comparator groups was identical to those described in preceding section 7.3.1 However, despite the evident differences in clinical phenotype there was no demonstrable difference in baseline gene expression in participants with different levels of disease activity after 3 months monotherapy (lowest adjusted p value = 0.50).

#### **7.4.3 Baseline Gene Expression Profile and Clinical Course – Mean DAS28 between 0 and 12 months**

The relationship between baseline expression profile and subsequent clinical course was examined to determine whether baseline expression profile patterns might have any ability to predict how patients fare over a period of time, regardless of treatment method. The presumption being, that particular patterns of treatment response (e.g. persistently active disease or persistently inactive disease) might be associated with distinct differences in underlying gene expression. Clinical course was characterised by calculating each participant's mean DAS28 between baseline and 12 months. Those participants with the highest mean DAS28 0-12 months would have experienced the most persistently active disease and, by extension the weakest treatment response, whilst those with the lowest mean DAS28 0-12 months would have experienced a much lesser overall disease burden and should be at a much lower risk of long term complications. The baseline expression profiles of the upper and lower quartile mean DAS28 0-12 months groups were compared. There were 18 participants in each group. In the lower quartile group mean DAS28 0-12 months ranged from 1.33 to 2.42 (median 1.93), whilst in the upper quartile group mean DAS28 0-12 months ranged from 3.72 to 5.67 (median 4.0). Mann Whitney U testing confirmed that the distribution of mean DAS28 0-12 months for each group were statistically distinct ( $p < 0.0001$ ). Baseline expression profiles did not appear to associate with particular patterns of clinical course since there were no differences in gene expression identified that were also associated with acceptable adjusted p values (lowest adjusted p value = 0.38).

#### **7.4.4 Dynamic Changes in Gene Expression Profile in Relationship to Changes in Disease Activity**

Neither gene expression nor inflammatory disease activity levels are static therefore the relationship between dynamic changes in gene expression and dynamic changes in disease activity measures was examined. Changes between baseline and 3 months for both gene expression and DAS28 were described by subtracting the baseline values from the 3 month values. The resulting values were then compared as continuous variables. This crude comparison did not identify evidence of differential gene expression (lowest adjusted p value = 0.43); however, the continuous nature of the comparator variables may have limited the comparison's ability to identify between group differences by not creating clearly separated phenotypic groups.

### **7.5 Additional Unplanned Phenotypic Groupings**

Sections 7.2 – 7.4 describe the outcomes of gene expression comparisons that were planned in advance of the laboratory analysis. However, the results so far do not suggest that there is any consistent evidence of differential gene expression within the peripheral blood of these clinically relevant phenotypic groupings. Potential reasons why this may be the case will be discussed in Section 7.6. Several comparisons (most notably groupings based upon gender) did identify differential gene expression between the comparator groups. Furthermore, following completion of the original analysis, the 3 month MSUS findings data became available for the subset of participants within the MSUS assessment group. The information provided by both of these developments was used separately to define several post hoc phenotypic groupings that underpinned additional comparisons of gene expression profiles

#### **7.5.1 Correcting for Influence of Gender on Gene Expression Profiles**

To correct for any potential skew that gender may have had on gene expression profiles in other comparisons a number of comparisons were rerun using only clinical outcome and gene expression data from female participants (n = 51, the larger gender group).

#### **Baseline Gene Expression Profile and Rheumatoid Factor Status – Females Only**

When the baseline gene expression profiles of rheumatoid factor positive (n = 31, 61%) and negative (n = 20, 39%) female participants were compared a single gene was differentially expressed (down regulated) using an adjusted p value threshold of 0.05. The microarray ID (7650093) was not associated with a recognised Gene ID and therefore it may be that this microarray location represents a control probe. Using an adjusted p value threshold of 0.15, one additional gene (CCDC28B), that had not been identified when the whole cohort was examined, appeared to be up-regulated in rheumatoid factor positive participants. The chromosomal locations and expression levels of these 2 genes are shown in Table 50

Illumina Gene Symbol	Gene Name	log Fold Change	Adjusted p Value	Chromosomal Location
7650093	?	-0.19	0.048	?
CCDC28B	Coiled-coil domain containing 28B	0.22	0.11	1p36.11-p34.2

**Table 50** - Genes which were differentially expressed in baseline samples of female participants grouped by rheumatoid factor status  
Adjusted p value threshold = 0.15

### **Baseline Gene Expression Profile and Anti-CCP Antibody Status – Females Only**

Twenty-six (51%) female participants were anti-CCP antibody positive and 25 (49%) were anti-CCP antibody negative. Separating female participants into groups based upon anti-CCP antibody status did not identify any significant differences in gene expression profile though did reduce slightly the adjusted p values of the genes with the highest rankings. Nevertheless, the lowest adjusted p value was still extremely high (0.94)

### **Baseline Gene Expression Profile and Erosion Status – Females Only**

Nineteen (37%) female participants exhibited plain x-ray erosions at presentation. Comparing baseline gene expression profiles between female participants with, and without, erosive x-rays also did not identify significant differential gene expression. Overall, adjusted p values were slightly lower than the whole cohort comparison, though did still remain extremely high (0.92).

### **Baseline Gene Expression Profile and Baseline Disease Activity Level – DAS28 Quartiles in Females Only**

Groups were formed based upon quartiles of the baseline DAS28 and between group gene expression comparisons were conducted between the upper and lower quartile groups. There were 13 participants in each group. The DAS28 range in the lower quartile group was 3.24 to 4.32 (median 3.66) and in the upper quartile group was 5.85 – 7.21 (median 6.47). Mann Whitney U testing confirmed that the distributions of DAS28 values in both groups were statistically distinct ( $P < 0.0001$ ). Even though there were clear differences in the measure of disease activity the adjust p values remained above any reasonable threshold that would have supported a significant between group difference in gene expression (lowest adjust p = 0.79).

### **7.5.2 Dynamic Change in Gene Expression Profile Between Baseline and 3 Months**

To determine, whether there was likely to be any dynamic change in gene expression profiles a paired analysis, comparing the baseline and 3 month expression profiles for all the participants was conducted. This analysis was not restricted by any phenotypic denominator or level of response. Overall, the whole cohort exhibited a significant improvement in DAS28. Median baseline DAS28 was 5.01, falling to 2.94 after 3 months ( $p < 0.0001$ ). Using an adjusted p value threshold of 0.05, paired comparisons of gene expression profile identified 19 genes whose expression patterns changed significantly between baseline and 3 months (Table 51). Five genes became upregulated and 14 genes became down regulated. All 19 differentially expressed

microarray locations were associated with previously identified human genes. Two genes that were down regulated were not listed as having a known chromosomal location (LOC651751 and LOC652493). Further, down regulated genes LOC647450 and LOC647506 were both associated with the same gene ID.

### **7.5.3 Incorporation of MSUS Disease Activity Assessment Findings into Phenotypic Groupings**

At low disease activity levels it is possible that the DAS28 will be insufficiently sensitive to identify those patients with truly quiescent disease since it has been consistently demonstrated that significant numbers of patients with DAS28 LDAS still exhibit evidence of synovitis on either clinical examination or imaging (121,192,200). Hence, the previously described comparisons (sections 7.3.2 - 7.3.4) may have been underpowered to identify differences in gene expression because the presumed LDAS / remission groups may still have contained some participants with active subclinical synovitis. With the availability of 3 month MSUS disease activity data it was possible to identify a further subgroup (n=5) of patients who had absolutely no ultrasonographic evidence of synovitis (total PD joint count = 0). This MSUS remission sub-group could be assumed to represent the very lowest level of inflammatory disease activity it was possible to identify using the available clinical and MSUS disease activity data. The three-month gene expression profiles of the MSUS remission sub-group were then compared to those of the 10 participants with the highest 3 month disease activity levels (who all had clinically evident synovitis). The DAS28 range in the MSUS remission sub-group was 1.049 to 3.581 (median 2.31) and 4.425 to 6.3 (median 5.0) in the high DAS28 sub-group. Mann Whitney U testing compared that the distribution of DAS28 values in both groups was statistically distinct (p=0.0007). The previously employed method of multiple t tests between both groups did not identify any significant differences in gene expression between the comparator groups (lowest adjusted p value = 0.54).

Illumina Gene Symbol	Gene Name	log Fold Change	Adjusted p Value	Chromosomal Location
<b>Up-regulated Genes</b>				
RNU1G2	U1 small nuclear 4	0.22	0.0088	1p36.1
RNU1-5	Variant U1 small nuclear 18	0.20	0.011	1q21.2
FCER1A	Fc fragment of IgE, high affinity I, receptor for alpha polypeptide	0.33	0.018	1q23
RNU1-3	U1 small nuclear 3	0.19	0.048	1p36.1
BCL11A	B-cell/lymphoma 11A (zinc finger protein)	0.13	0.057	2p16.1
<b>Down-regulated Genes</b>				
IGJ	Immunoglobulin J polypeptide	-0.53	0.016	4q21
MGC29506	Marginal zone B and B1 cell-specific protein	-0.29	0.018	5q31.2
SMN1	Survival of motor neuron 1	0.088	0.044	5q13.2
LOC653618	Similar to hypothetical protein LOC349196	-0.093	0.0022	8p23.1
TXNDC11	Thioredoxin domain containing 11	-0.11	0.018	16p13.13
TAF15	TAF15 RNA polymerase II	0.28	0.00042	17q11.1-q11.2
C17ORF79	Chromosome 17 open reading frame	0.20	0.0055	Chr 17
IGLL1	Immunoglobulin lambda-like polypeptide 1	-0.11	0.016	22q11.23
MIAT	Myocardial infarction associated transcript	-0.66	0.0004	22q12.1
LOC651751	Similar to Ig kappa chain V-II region RPMI 6410 precursor	-0.20	0.0044	Not known
LOC647450	Similar to Ig kappa chain V-I region HK101 precursor	-0.39	0.0022	Chr 2
LOC652493	Ig kappa chain V-I region HK102-like	-0.54	0.0032	Not known
LOC647506	Similar to Ig kappa chain V-I region HK101 precursor	-0.58	0.0036	Chr 2
LOC642113	Ig kappa chain V-III region HAH-like	-0.45	0.0053	Chr 2
LOC652694	Similar to Ig kappa chain V-I region HK102 precursor	-0.53	0.0055	Chr 2

**Table 51** – Genes which underwent changes of expression between baseline and 3 months  
Adjust p value threshold = 0.05

### 7.5.3 Comparison of Differences Between Baseline and 3 Month Gene Expression Profiles in Patients Who Achieved DAS28 Remission After 3 Months

Changes over time of microarray analyses are often correlated to corresponding changes in a designated disease state and are interpreted as being causally associated with the change in the disease state under investigation. However, peripheral blood gene expression profiles do not simply relate to a single disease state and may be influenced by multiple additional intrinsic factors; such as, environmental exposures, drug therapy, intercurrent illness, co-morbidities and

genetic predisposition. At a patient level, many of these potential influences will remain constant; however, at a group level they may not be equally distributed amongst all the patients. To minimise the influence of patient related factors on the observed gene expression profiles, the baseline and 3 month gene expression profiles of participants who achieved DAS28 remission after 3 months were compared. Participants required at least moderate disease activity to enter the study; therefore, those patients who achieved DAS28 remission after 3 months were likely to have undergone a profound clinical change which should increase the likelihood of observing a corresponding dynamic change in gene expression. Results from 34 participants were included in the analysis and a clear clinical response to 3 months of DMARD monotherapy was observed. Median DAS28 score fell from 4.40 (IQR 3.71 – 5.44) to 2.00 (IQR 1.59 – 2.30) and Wilcoxon Signed-Ranks test confirmed a statistically significant change between the two time points ( $p < 0.0001$ ). However, despite the clear change in clinical state, there were no statistically significant differences in gene expression profile evident between the 2 time points (lowest adjusted p value = 0.99). Repeating the comparison using more widely spaced samples (e.g baseline and 18 months) may allow more time for shifts in gene expression to become apparent and may determine whether any lead bias was present

A similar comparison was conducted using the subset of 14 participants whose MSUS disease activity assessment at 3 months identified PD signal in one or no joints since, under the proposed MSUS definitions of disease activity, this state effectively equates to MSUS LDAS and/or remission and was not felt to justify DMARD escalation. Since at baseline all participants had exhibited evidence of active synovitis (either clinically and/or ultrasonographically) those who exhibited a MSUS PD joint count of 1 or zero after 3 months were presumed to have experienced the biggest swing in overall inflammatory disease activity. Of the participants who underwent MSUS disease activity assessment after 3 months follow-up, 9 had a PD joint count of 1 and 5 had no evidence of PD signal in any joint. Of these 14 fourteen patients, 6 (43%) also fulfilled the ACR-Boolean definition for remission. Overall, this subgroup of participants had experienced a good response to initial DMARD monotherapy as measured by DAS28; median DAS28 at baseline was 4.79 (IQR 4.34-5.45) falling to 2.34 (IQR 2.05-2.99) after 3 months. The median fall in DAS28 was 2.29 (IQR 2.29) and a Wilcoxon Signed Rank comparison confirmed that there had been a significant change in the distribution of DAS28 values between the two time points ( $p = 0.001$ ). Despite the clear improvements in clinical and MSUS disease activity assessments there was no evidence of differential gene expression between the two sample time points (lowest adjusted p value = 0.99)

#### **7.5.4 Relationship of Interferon Response Gene Expression to Disease Activity Levels**

Whilst conducting this study, two linked papers were published describing how expression of interferon response genes changed in response to rituximab therapy and suggesting that elevated pre-treatment interferon response gene levels were predictive of treatment non-response (355,356). Both papers attempted to manipulate gene expression data, to produce a single, numerical score that was both objective and easily understood. In the first paper, the 3

month mean expression level of 6 type I interferon response genes (IRG) (RSAD2, IFI44, IFI44L, HERC5, LY6E, Mx1) increased in patients who experienced a good 6 month response to rituximab (356). In the second paper, an elevated pre-treatment mean expression level of 8 type I IRG (LY6E, HERC5, IFI44L, ISG15, MxA, MxB, EPSTI1, RSAD2) was predictive of non-response to rituximab after 6 months (355). Using the available clinical outcome and expression profile datasets from this research, an analysis was conducted to determine whether the IRG signatures either 1. exhibited the same changes of expression in response to DMARD therapy and/or 2. were predictive of response to DMARD monotherapy after 3 months and overall response to step-up DMARD therapy after 12 months.

## **Analysis 1 – relationship between change in IRG score and disease activity following commencement of DMARD therapy (after Vosslander et al, 2011 (356))**

### *Methods*

1. Comparisons groups defined based upon DAS28 score
  - i. After 3 months –  $\Delta$ DAS28 < 1.2 vs  $\Delta$ DAS28  $\geq$  1.2
  - ii. After 6 months -  $\Delta$ DAS28 < 1.2 vs  $\Delta$ DAS28  $\geq$  1.2
  - iii. After 12 months -  $\Delta$ DAS28 < 1.2 vs  $\Delta$ DAS28  $\geq$  1.2
  - iv. After 3 months – DAS28 < 3.2 vs DAS28  $\geq$  3.2
  - v. After 6 months - DAS28 < 3.2 vs DAS28  $\geq$  3.2
  - vi. After 12 months - DAS28 < 3.2 vs DAS28  $\geq$  3.2
2. Calculation of IRG score using baseline and month 3 gene expression data for each participant  
IRG score = mean of expression levels of RSAD2, IFI44, IFI44L, HERC5, LY6E, Mx1
3. Calculation of IRG ratio (IRG T3:T0) between month 3 and baseline mean IRG score for each participant
4. Between group comparison of pooled IRG T3:T0 values: participants segregated by treatment response comparator groups. Between group comparisons conducted using Mann-Whitney-U test and represented graphically using boxplot of median, interquartile range, minimum and maximum values

### *Results*

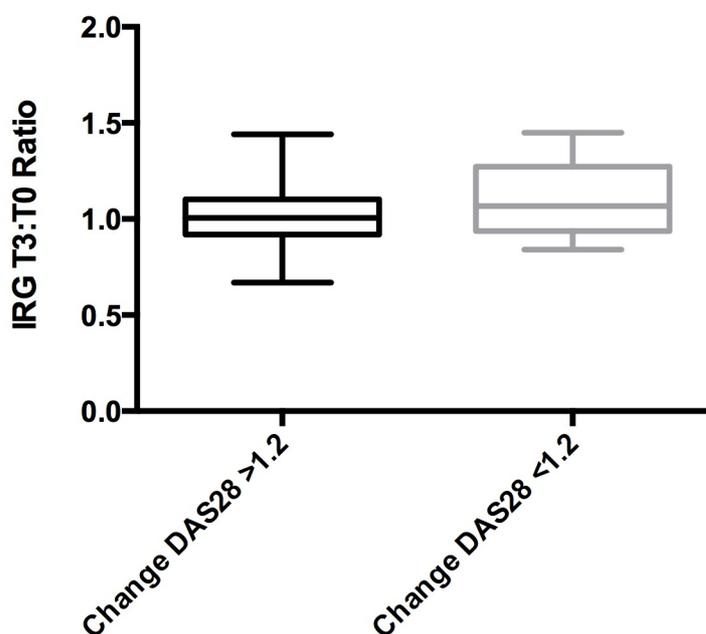
Paired sets of gene expression and DAS28 data were available for 74 participants between baseline and follow-up month 3 and between baseline and follow-up month 6 and for 69 participants between baseline and follow-up month 12. Overall, the dynamic change in the mean IRG score between baseline and 3 months was not associated with either a significant improvement in DAS28, nor likelihood of attaining low disease activity, after 3, 6 or 12 months of follow-up. The summary statistics suggest that the majority of 3 month IRG T3:T0 ratio values were closely associated with the value 1.0 (mean 1.0,  $\pm$ SD 0.18) implying that for the majority of

patients there was very little change in the overall expression levels of IRG genes between baseline and month 3.

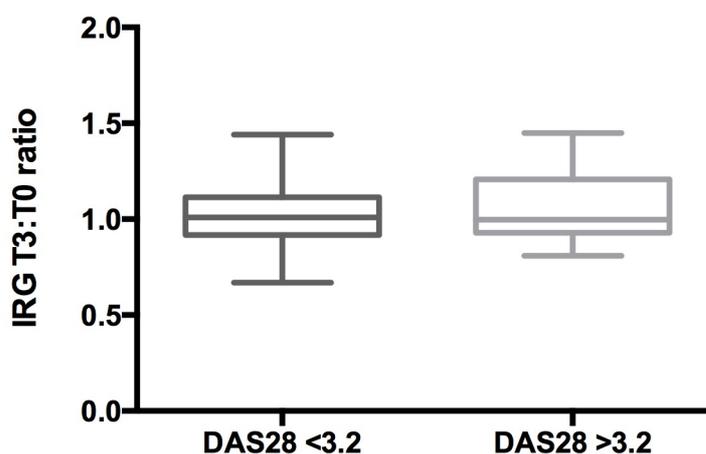
### 1. Response to DMARD monotherapy after 3 months

After 3 months of DMARD monotherapy, 61 participants were classified as responders and 13 participants were classified as non-responders. For responders, the median change in DAS28 was -2.27 and for non-responders the median change was 0.12; there was a statistically significant difference in the degree of change between the groups (Mann-Whitney  $p < 0.0001$ ). However, the IRG T3:T0 ratios for both groups were very similar (Graph 43). Between group comparisons showed that there was no statistically significant difference between the groups (Mann-Whitney  $p = 0.16$ ). A correlation analysis comparing the 3 month IRG T3:T0 ratio to the corresponding change in DAS28 suggested a positive correlation between the values (Pearson's co-efficient:  $r = 0.33$ , 95%CI 0.11-0.52,  $p = 0.0045$ )

After 3 months, 51 participants exhibited a DAS28 score  $< 3.2$  (low disease activity) and 23 participants exhibited a DAS28 of 3.2 or higher (moderate disease activity above); between group comparisons suggested that the distribution of DAS28 scores were statistically distinct for each group (median DAS28 2.3 vs 4.2 respectively,  $p < 0.0001$ ). The distribution of IRG T3:T0 ratios for both these groups were statistically similar (Graph 44). The median IRG T3:T0 ratio for both groups was 1.0 (Mann Whitney,  $p = 0.68$ )



**Graph 43** – IRG T3:T0 ratio according to 3 month clinical response. Clinical response defined as change in DAS28  $> 1.2$

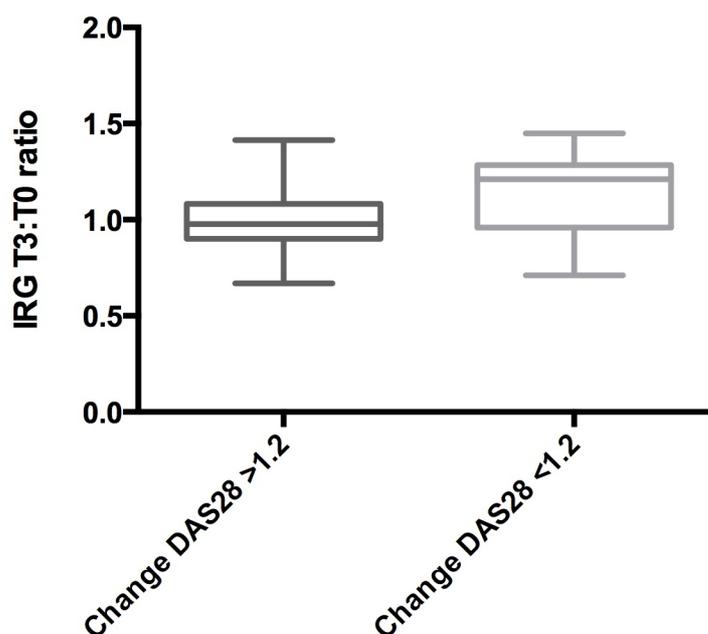


**Graph 44** – IRG T3:T0 ratio between patients with low disease activity (DAS28 <3.2) and moderate disease activity or higher (DAS28 >3.2) at 3 months

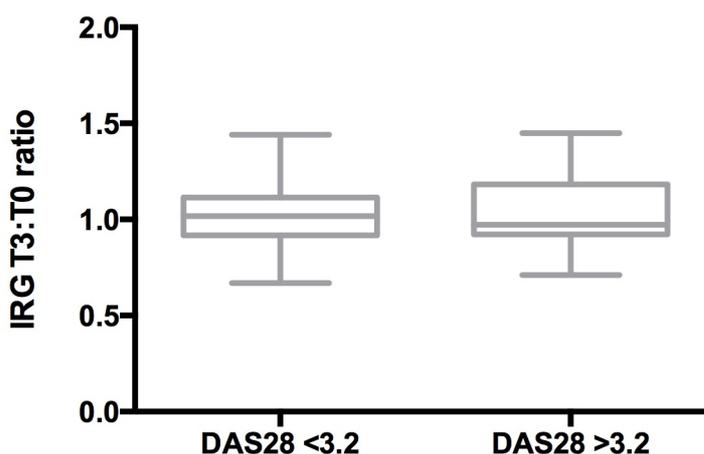
2. Response to step-up DMARD therapy after 6 months

This comparison most closely resembled the analysis described by Vosslamber et al (356). After 6 months of treatment, 56 participants were classified as responders and 18 were classified as non-responders. In the responder group, the median fall in DAS28 was -2.4 and in the non-responder group the median fall in DAS28 was -0.19; the between group comparison confirmed a statistically significant difference in the mean fall in DAS28 between both groups (Mann Whitney  $p < 0.0001$ ). Whilst there was considerable overlap in the distribution of values, between group comparisons did suggest that there was a statistically significant difference between the IRG T3:T0 ratios for both groups (Graph 45). However, the pattern observed for this research was contrary to that described in the original paper. Participants defined as non-responders tended to have higher median IRG T3:T0 ratios than responders (median IRG T3:T0 ratio: responders 1.0 vs non-responders 1.2,  $p = 0.0046$ ). Correlation analysis identified a weak positive correlation between 3 month IRG T3:T0 ratio value and change in DAS28 between baseline and 6 months (Pearson's co-efficient:  $r = 0.29$ , 95%CI 0.021-0.45,  $p = 0.033$ ).

After 6 months of DMARD therapy, 43 participants had attained a DAS28 LDAS (median DAS28 = 2.1), whilst 31 participants continued to exhibit a DAS28 of 3.2 or greater (median DAS28 = 4.1). The distribution of DAS28 values between both groups was significantly different (Mann Whitney  $p < 0.0001$ ). However, there was very little difference evident in the 3 month IRG T3:T0 ratios between the groups (Graph 46). The median IRG T3:T0 ratio for the DAS28 <3.2 group was 1.0 and the median IRG T3:T0 ratio for the DAS28 >3.2 or greater group was 0.97; Mann Whitney test between the groups returned a  $p$  value of 0.97. Therefore, no association between 3 month IRG T3:T0 ratio and 6 month disease activity level was identified in this comparison



**Graph 45** – Association between 3 month IRG T3:T0 ratio and clinical response at 6 months. Clinical response defined as a fall in DAS28 of 1.2 or greater

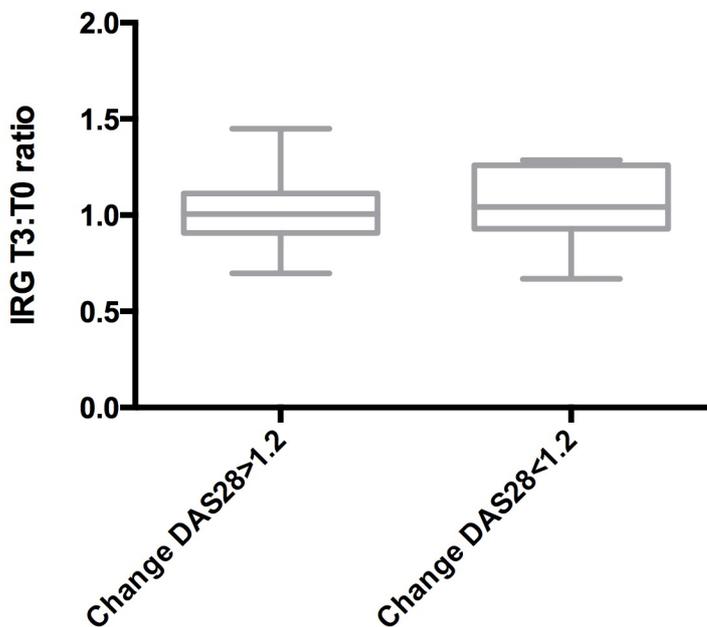


**Graph 46** – IRG T3:T0 ratio between patients with low disease activity (DAS28 <3.2) or moderate disease activity or higher (DAS28 >3.2) at 6 months

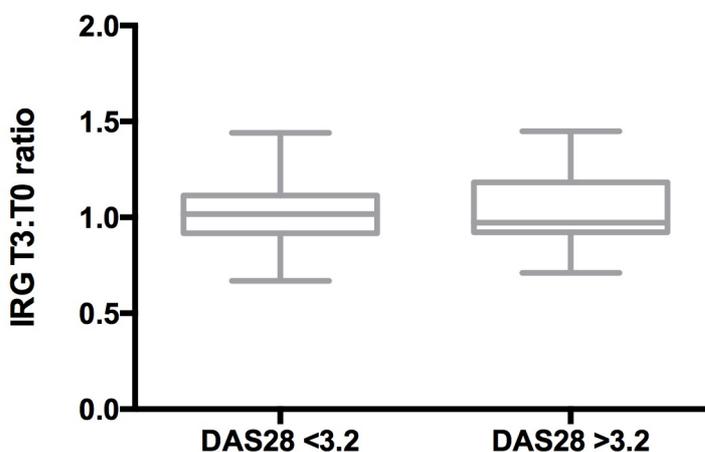
### 3. Response to step-up DMARD therapy after 12 months

After 12 months of step-up DMARD therapy, 57 participants were classified as responders (median change in DAS28 = -2.6) and 12 were classified as non-responders (median change in DAS28 = -0.60). Mann Whitney testing confirmed that there was a statistically significant difference in the degree of change in DAS28 between the groups ( $p < 0.0001$ ). However, there was no statistically significant difference evident in the 3 month IRG T3:T0 ratio between either group (Graph 47). Median IRG T3:T0 ratio for both groups was 1.0 and between group comparisons returned a non-significant p value (Mann-Whitney  $p = 0.62$ ). Correlation analysis suggested that there was a weak positive correlation between 3 month IRG T3:T0 ratio and change in DAS28 between baseline and 12 months (Pearson's coefficient:  $r = 0.24$ , 95%CI 0.00053-0.44,  $p = 0.050$ )

After 12 months of follow-up, 48 participants had attained a DAS28 <3.2 and 21 participants exhibited a DAS28 > 3.2. Once again, the distribution of DAS28 scores between both groups was statistically distinct (median DAS28 1.8 vs 4.0 respectively,  $p < 0.0001$ ), though there was no association evident between 3 month IRG T3:T0 ratio and disease activity state at 12 months (Graph 48). The median 3 month IRG T3:T0 ratio for both groups was 1.0 and between group comparisons returned a non-significant p value (Mann Whitney  $p = 0.70$ ).



**Graph 47** – Association between 3 month IRG T3:T0 ratio and clinical response at 12 months. Clinical response defined as change in DAS28 of 1.2 or greater



**Graph 48** - IRG T3:T0 ratio between patients with low disease activity (DAS28 <3.2) or moderate disease activity or higher (DAS28 >3.2) at 12 months

## **Analysis 2 – relationship between baseline IRG score and clinical response following commencement of DMARD therapy (after Raterman et al, 2012 (355))**

### *Methods*

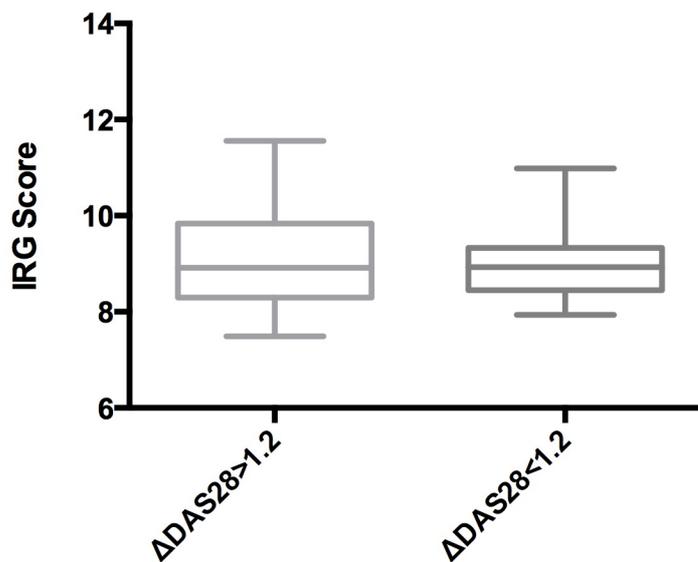
1. Comparison groups defined based upon DAS28 score
  - i. After 3 months –  $\Delta$ DAS28 < 1.2 vs  $\Delta$ DAS28  $\geq$  1.2
  - ii. After 12 months -  $\Delta$ DAS28 < 1.2 vs  $\Delta$ DAS28  $\geq$  1.2
  - iii. After 3 months – DAS28 < 3.2 vs DAS28  $\geq$  3.2
  - iv. After 12 months - DAS28 < 3.2 vs DAS28  $\geq$  3.2
2. Calculation of IRG score using baseline expression data for each participant
  - IRG score = mean of expression levels of LY6E, HERC5, IFI44L, ISG15, MxA, MxB, EPSTI1 and RSAD2
3. Between group comparison of pooled IRG scores; participants segregated by treatment response comparator groups. Between group comparisons conducted using Mann-Whitney-U test and represented using boxplot of median, interquartile range, minimum and maximum values

### *Results*

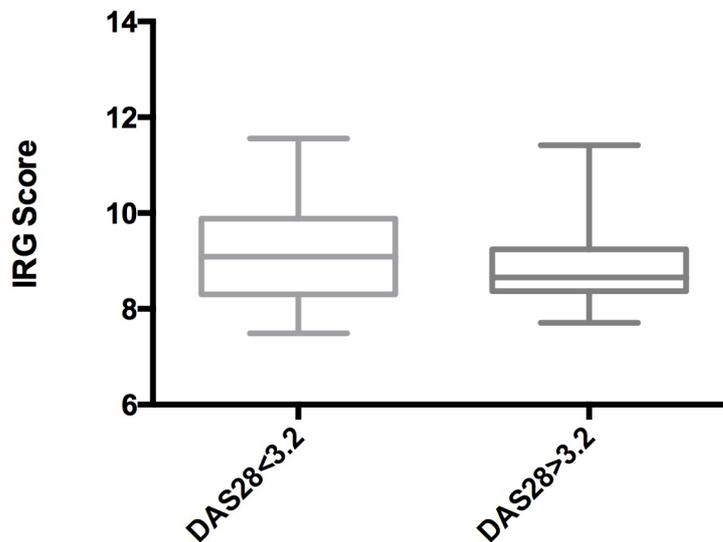
Overall, no significant differences in baseline IRG scores were observed when patients were stratified by either change in DAS28 or absolute DAS28 value at either 3 months or 12 months.

1. Response to DMARD monotherapy after 3 months.

Using both definitions of treatment response created comparator groups with statistically different levels of measured treatment response. After 3 months of DMARD monotherapy, 61 participants were classified as responders and 13 participants were classified as non-responders (median change DAS28: -2.77 vs +0.12,  $p < 0.0001$ ). However, the median and distribution of mean IRG scores was very similar for both groups (median IRG score: 8.91 vs 8.93,  $p = 0.88$ ) (Graph 49). Similarly, 51 participants exhibited a 3 month DAS28 less than 3.2 and 23 exhibited a DAS28 of 3.2 or higher (median DAS28: 2.34 vs 4.18,  $p < 0.0001$ ). However, despite the differences in disease activity levels, there was no statistically significant difference evident in the distribution of IRG scores between the groups (median IRG score: 9.09 vs 8.66,  $p = 0.45$ ) (Graph 50)



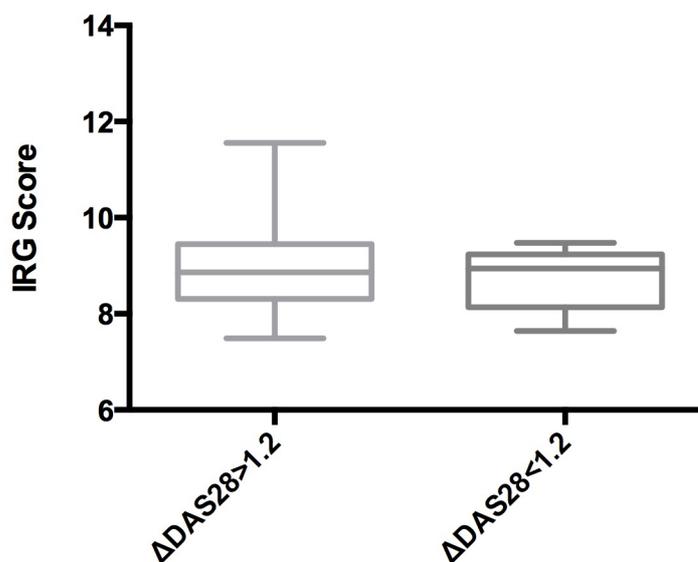
**Graph 49** – Baseline IRG score in relation to 3 month clinical response  
Clinical response defined by change in DAS28 from baseline



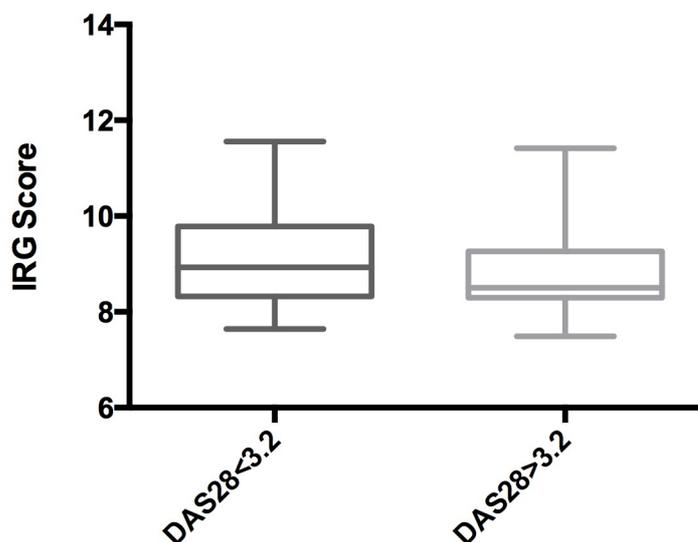
**Graph 50** – Baseline IRG score in relation to 3 month clinical response  
Clinical response defined by value of DAS28 at 3 months

2. Response to step-up DMARD therapy after 12 months

After 12 months follow-up, 57 participants were classified as responders and 12 were classified as non-responders (median change DAS28: -2.64 vs -0.60,  $p < 0.0001$ ). Both groups exhibited a very similar distribution in the change of IRG after 12 months (median IRG score: 8.86 vs 8.94,  $p = 0.63$ ) though the total range for the positive response group was higher (Graph 51). Forty eight participants attained a DAS28 less than 3.2 and 21 participants didn't (median DAS28: 1.79 vs 3.96,  $p < 0.0001$ ). However, despite the differences in clinical response, there was no statistically significant difference identified in the associated distribution of IRG scores (median IRG score: 8.93 vs 8.50,  $p = 0.34$ ) (Graph 52).



**Graph 51** – Baseline IRG score in relation to 12 month clinical response  
Clinical response defined by change in DAS28 from baseline



**Graph 52** – Baseline IRG score in relation to 12 month clinical response  
Clinical response defined by value of DAS28 at 12 months

## 7.6 Discussion

Previous studies describing the relationship between gene expression and RA have tended to focus on either disease pathogenesis and/or a limited aspect of disease phenotype. So far, no studies have attempted a comprehensive phenotypic / genotypic approach such as that described herein. As a complex disease, whose clinical course is often heterogenous and dynamic, RA phenotype can be considered from a number of different angles: 1. as a single fixed factor at a single time point (e.g. RhF status, anti-CCP antibody status); 2. as a static measure of disease activity at a single time point (e.g. DAS28), 3. as a dynamic change in a measure between time points and 4. as the overall level of disease activity over a given time period (e.g. mean DAS28). Therefore, the relationship between gene expression profiles and RA phenotype can also be considered from a variety of angles: 1. the relationship between gene expression profile and RA phenotypic factor at a single time point; 2. the relationship between gene expression patterns at a

single time point and a future RA phenotype (i.e. the predictive ability) and 3. how well dynamic changes in RA phenotype (most commonly a measure of disease activity) are reflected in corresponding changes in gene expression.

Taken together the following conclusions can be drawn from the analyses:

#### *Baseline gene expression profile and baseline disease phenotype*

1. Participant gender was associated with the highest number of differentially expressed genes. In females, 25 genes were down-regulated and 36 genes were up-regulated in relation to males. Accounting for gender did not identify any new evidence of differential gene expression between the phenotypic comparator groups
2. Five genes were up-regulated in RhF positive patients compared to RhF negative patients
3. Current tobacco smoking was associated with up-regulation of a single gene (LRRN3) in relation to former smokers and non-smokers. In former smokers, one gene (HIATL1) was up-regulated in relation to non smokers.

#### *Gene expression profiles in relation to disease activity*

4. There was no evidence of differential gene expression between upper and lower quartile groups of DAS28 at either baseline or follow-up month 3; despite the groups appearing phenotypically different at both time points
5. The expression pattern of 19 genes changed following three months of DMARD monotherapy. Five genes became up-regulated and 14 genes became down-regulated. However, the subgroups of participants who experienced the biggest improvement in disease activity did not exhibit evidence of changing gene expression patterns
6. There was no evidence of differential gene expression between the subgroups of patients with the greatest phenotypic difference in disease activity after 3 months of DMARD therapy. In particular, there was no evidence of differential gene expression between MSUS-defined remission and patients with clinically active disease

#### *Baseline gene expression as a predictor of future disease activity*

7. At baseline, three genes were up-regulated in participants who subsequently qualified for treatment escalation after 3 months of DMARD monotherapy
8. Baseline gene expression patterns did not appear predictive of disease activity status after 3 months of DMARD monotherapy nor clinical course over 12 months
9. Mean baseline interferon response gene score neither predicted treatment response nor disease activity state after 3 months and 12 months of treatment

### **7.6.1 Relationship between baseline phenotypic characteristics and gene expression profile**

There has been very little work that has systematically considered differences in peripheral blood gene expression based upon RA-associated phenotypic factors. In this study, participant's gender, rheumatoid factor status and smoking status were associated with differential expression

of small numbers of genes. Conversely, anti-CCP antibody status and baseline erosive status did not appear to be associated with any significant differences in gene expression.

### *Gender*

Unsurprisingly, segregating participants by gender identified the highest number of genes with differential expression between the groups. Sixty six genes exhibited differential expression; 36 were up-regulated in females and 25 were down regulated. All 14 X chromosome related genes were up-regulated in females and all 16 Y chromosome related genes were down-regulated. This pattern is partly explained by the analysis method, since gene expression levels of male participants were compared relative to female participants. The pattern identified for differential gene expression between gender groups bore similarities to the patterns described in other studies of gender associated differences in gene expression. Four of the X chromosome related genes (EIF1AX, EIF2S3, RPS4X, XIST) and 4 of the Y chromosome related genes (RPS4Y1, EIF1AY, CYorf15B, USP9Y) had previously been shown to be differentially expressed in a comparison between healthy gender groups (357,358); though that data had been analysed using the Significance Analysis of Microarrays technique, and a slightly higher adjusted p value threshold (0.065). Four of the 66 differentially expressed genes identified by this study (XIST, EIF1AX, EIF2S3, RPS4X) were included within the 19 genes exhibiting gender related differential expression in peripheral blood from another study conducted on normal subjects (359-361). Similarly, a recent study by Xu et al, which used a stringent adjusted p value threshold ( $p < 0.01$ ), identified 105 differentially expressed genes between gender groups (362). Interestingly, all of the top 10 gender-associated genes were sex chromosome related and 8 of these were also identified as being differentially expressed by this study (XIST, RPS4Y1, EIF1AY, CYorf15A, CYorf15B, USP9Y, UTY, PRKY). Using an adjusted p value threshold of 0.01, eleven of the 37 differentially expressed genes identified by this study (CA5B, CYorf15A, CYorf15B, EIF1AY, HDHD1A, PRKX, PRKY, RPS4Y1, UTY, XIST, ZFY) were also identified within the 104 differentially expressed genes identified by Xu et al. Further, the expression patterns of sex chromosome related marker genes (XIST and RPS4Y1) for this study were similar to those reported by Wu et al (i.e. XIST was up-regulated in females whilst RPS4Y1 was up-regulated in males).

These findings are similar, but not identical, to those of previous studies in healthy subjects. In fact, the results reported herein seem to be the first reported description of gender related gene expression differences in RA. It seems likely that the majority of observed differences in gene expression are gender rather than RA related. It also seems unlikely that differences in gender related gene expression will have had a sufficient impact on overall gene expression to confound the observations of other RA-related comparisons; especially since the 66 gender-related genes comprise a very small proportion of the 47,000 gene probes on the Illumina HumanHT Beadchip. Nevertheless, additional analyses were conducted using data from female participants in an attempt to account for the potential influence of gender on gene expression when participants were grouped according to RA-related characteristics. In fact, focusing on female participants alone did not identify any new differences in gene expression and also altered the previous observation that RhF status was associated with some differential gene expression. Whilst this

suggests that the inclusion of males in the original analysis may have influenced the observed gene expression profiles, it was not related to an imbalance in the proportion of males in either group since there was no statistical difference in sex distribution between the RhF status groups (Percentage males - RhF +ve 37% vs RhF -ve 29%,  $p=0.62$ ). Similar comparisons conducted in female participants grouped according to their anti-CCP antibody status, erosive status of baseline DAS28 disease activity did not identify any significant between-group differences of gene expression.

#### *Tobacco smoking status*

Several previous studies have shown that tobacco smoking affects gene expression systemically (353,363) and within the small airways (354). Further, the relationship between tobacco smoking and increased risk of developing rheumatoid arthritis, particularly RhF and anti-CCP antibody positive RA, is now well established (364) and may in part be initiated by smoking-induced activation of citrullination within the respiratory epithelium (97). Given the known complexities of RA pathogenesis and the increasing recognition of its link to tobacco smoke exposure it is surprising that this research did not identify a greater number of differentially expressed genes. The differential expression of the LRRN3 gene between current smokers and non-smokers and current smokers and former smokers replicates a previously identified association (353) and is therefore more likely due to the influence of tobacco smoke exposure on whole blood gene expression generally, rather than RA per se. It seems reasonable to presume that RA patients whose disease development has been influenced by tobacco smoke exposure might feasibly exhibit differences in gene expression that are not present in RA patients with a different aetiological background. However, the cohort investigated by this study may have been too small and prone to the influence of confounders to accurately reflect the complexities of the interactions between genetic predisposition, tobacco smoke exposure and RA pathogenesis. Furthermore, it is possible that the gene expression changes induced by tobacco smoke exposure might be better demonstrated in alternative target tissues (such as synovium or respiratory epithelium) rather than whole peripheral blood which is more exposed to the influence of external factors and co-existing illness.

#### *Rheumatoid factor status*

From this study, an adjusted p value threshold of 0.05 identified 2 genes (SRI, NAIP) that were up-regulated in RhF positive participants compared to RhF negative patients. Increasing the adjusted p value threshold to 0.15 identified a further 4 up-regulated genes (LOC648984, LOC728519, AP1S1, BCS1L) in RhF positive participants. Interestingly, the same genes were not found to be differentially expressed when the comparison was restricted to female participants. These initial findings do suggest that RhF status might be associated with a specific pattern of gene activation. However, it is worth re-iterating that the gene list was not apparent until the adjusted p value threshold had been arbitrarily raised. Equally, it is well recognised that RhF expression is not specific to RA. Several other conditions (e.g. chronic hepatitis C infection, Sjogrens syndrome, systemic lupus erythematosus) have also been associated with RhF positivity and approximately 11% of healthy subjects may also exhibit rheumatoid factors without having RA (365,366). Therefore, even if the gene signature tentatively reported by this research

were to be confirmed by qRT-PCR, and in an independent validation cohort, its relevance to RA pathogenesis would need to be carefully explored.

Very few studies have so far described the relationship between gene expression profile and RhF status. A small study of 8 RhF positive and 6 RhF negative RA patients suggested that there were no differences in PBMC gene expression between the groups nor between sub-groups of patients with the three highest and three lowest RhF titres (367). Van Baarsen et al described gene expression profiles in auto-antibody positive arthralgic patients who had not yet developed frank RA (368). Within a subset of 9 RhF positive patients, a small (unspecified) number of genes were differentially expressed compared to 6 healthy control subjects. Interestingly, within a larger cohort of 109 auto-antibody positive arthralgic patients, the gene expression profiles of those patients who subsequently progressed to clinical inflammatory arthritis tended to cluster with the gene expression profiles of patients with established RA.

#### *Anti-CCP antibody status*

Despite a positive anti-CCP antibody status being highly specific, and predictive, for RA, this analysis did not identify any differential gene expression between anti-CCP antibody positive and anti-CCP antibody negative participants. This is particularly surprising since the development of ACPA antibodies is at least partly genetically determined (90,322). Further, it seems feasible to presume that differences in whole blood (i.e. systemic) gene expression might arise from the interaction between environmental (i.e. smoking) and genetic (HLA-DRB4, PTPN22) factors that culminate in the expression of ACPA antibodies. It is possible that the characteristics of the anti-CCP antibody comparator groups may have confounded the gene expression analysis. Whilst the groups were similarly sized (anti-CCP positive 41, anti-CCP negative 36) and balanced for gender distribution (%females: anti-CCP positive 69%, anti-CCP negative 63%), the anti-CCP positive group unsurprisingly contained a significantly higher number of current smokers (anti-CCP positive 39% vs anti-CCP negative 17%,  $p=0.045$ ). Given existing evidence that tobacco smoking does influence gene expression in peripheral blood (353,363) it is possible that the gene expression influence of tobacco smoke may have modulated the gene expression changes observed in relation to anti-CCP status. Nevertheless, correcting for tobacco smoking status between the anti-CCP comparator groups and rerunning the gene expression comparison did not identify any new differentially expressed genes. Further, it is possible that the anti-CCP antibody comparator group was too small to account for all the potential confounding factors that might also influence gene expression and ACPA antibody expression (e.g. smoking status, gender and hormone status, weight, alcohol intake, presence of periodontal disease, HLA-DRB4 and PTPN22 genotype). It is also important to consider that a significant subset of RA patients do not develop ACPA antibodies but do still develop a clinical and histological expression of synovitis that is very similar to ACPA positive RA patients (369). Therefore, if the eventual histological and clinical expression phenotypes of ACPA positive and negative RA patients are similar (i.e. a common pathway) it is perhaps less surprising that there were no differences in whole blood gene expression identified either. It is possible that ACPA expression may be associated with differential gene expression within the target tissues (such as respiratory epithelium or synovial tissue); however, this requires further investigation and targeted tissue sampling

Several studies have used gene expression data to perform hierarchical clustering exercises that cluster patients into distinct sub-groups based upon gene activation state and pathway analysis (for example: van der Pouw Kraan et al 2003 (229)). Their findings need to be carefully interpreted, and often inferred, since in many cases the original studies were not designed to investigate the relationship between RA phenotype and gene expression profile and any reported findings in relation to ACPA status are often supplementary. Taken together the available results of previous studies provide a signal that ACPA status might be associated with detectable differences in gene expression. Junta et al reported differential expression of 101 genes between anti-CCP positive and negative RA patients (370), van Baarsen et al reported that RA patients with high levels of synovial inflammation exhibited different patterns of gene expression within synovial tissue and also tended to have higher ACPA titres than patients with low levels of synovial inflammation (371). Further, van Baarsen et al have also reported that auto-antibody patients who develop RA have similar gene expression profiles to patients with established RA (i.e. they can be differentiated from auto-antibody positive patients who do not develop RA); though they do not specifically describe whether there are differences between different subsets of autoantibody patients nor whether gene expression in auto-antibody negative RA patients differs at all from auto-antibody positive RA patients (368). The combined findings of these studies is not conclusive. Furthermore, due to relatively small sample sizes, heterogeneity of design and study population and the potential for external confounders their clinical relevance is uncertain. Their results are also directly contrary to those reported herein. Even though this study was conducted on a notably larger population than any of the previously reported studies it still did not identify any evidence of differential gene expression between anti-CCP positive and negative participants. This could be partly explained by the cohort being too small to minimise the influence of all other confounders and by not having optimised the statistical analysis sufficiently to account for external confounders.

#### *Baseline x-ray status*

The emergence of radiographic erosions is an important step during the time course of RA. Whether certain patients are inherently predisposed to the development of erosions because of pathogenetic characteristics of their RA, or whether the development of erosions represents a point where the burden of joint destruction exceeds the bone's ability to repair remains unclear. Either way, since the development of erosions has long-term prognostic implications it seemed prudent to determine whether their presence at presentation might be associated with differences in gene expression. Overall, the available results suggest that the presence of plain x-ray erosions at presentation is not associated with evidence of underlying differences in gene expression. When considering these findings it is important to consider the characteristics of the grouping variable used to construct the comparator groups. Since the formal grading of the radiology outcomes has not yet occurred the baseline plain x-ray images have not yet been systematically graded for the presence of erosions. Rather, in order to allow a preliminary analysis, participants were grouped depending on whether the routine reporting (but not grading) by NHSGGC staff radiologists described the presence of erosions. All plain x-ray images collected within NHSGGC radiography departments undergo verbal reporting by staff radiologists

with a large variety of experience in musculoskeletal radiography and this process may have led to inconsistencies in how baseline x-rays were categorised. Once the formal grading of the plain x-ray images has been completed the comparison will be rerun with participants categorised based upon the components of the van-der-Heijde/Sharp score. An additional comparison is also planned to determine whether future deterioration in plain x-ray appearances is in anyway associated with baseline gene expression profile.

A single study has already reported on the association between erosive plain x-ray damage and differences in baseline PBMC gene expression profile in 96 African-American early RA patients (372). Using a relatively generous adjusted p value threshold (0.30), 1138 genes were shown to exhibit differential expression between patients with mild and severe erosive change on plain x-rays of hands and feet. There was far less striking evidence of differential gene expression between patients who did, and did not, exhibit progressive radiographic damage.

In fact it may not be possible to identify a single gene expression pattern that associates strongly with either the presence or future development of erosions since their development may instead represent the end-outcome of several overlapping pathogenetic processes associated with the presence of inflammatory synovitis (inflammatory cell activation and migration, pannus formation and proliferation, cytokine and osteolytic enzyme release, synovial fluid exudation, angiogenesis, apoptosis) (373). The activation states of each of these different processes may themselves be associated with differential gene expression; therefore, it may be that the gene expression profile for each process must be first described and correlated with the profiles of all other associated processes before a specific gene expression profile relating to the presence or development of plain x-ray erosions can be proposed.

### **7.6.2 Relationship between RA disease activity state and gene expression profile**

The clinical phenotype of RA patients with persistently active disease is very different from those with persistent LDAS or remission, even if the presenting substrates have been similar. Therefore it was hypothesised that differences in disease activity state, representing different stages of immune activation, might also be associated with differences in underlying gene expression. Whilst differences in gene expression might provide interesting insights into the activity of underlying inflammatory pathways for the findings to be clinically useful they must also relate to recognisable clinical scenarios. Therefore participants were formed into comparator groups based upon permutations of the DAS28 score, a widely used composite measure with well recognised thresholds for different disease activity states. At presentation, most participants had not yet commenced immunomodulatory therapy and the distribution of their DAS28 scores was skewed towards moderate-high disease activity. Likewise after 3 months follow-up, most participants were starting to experience the benefits of commencing DMARD therapy, so the distribution of their DAS28 score was skewed towards moderate-low disease activity. Therefore, in order to include the whole spectrum of disease activity comparisons were performed using both the baseline and 3 month data. A comparison combining the 2 datasets, to encompass the whole range of possible

disease activity states, was not performed since most participants would be represented twice within a single dataset and statistical corrections for multiple inclusions of some participants would become exceedingly complicated.

The DAS28 provides a continuous measure of disease activity; therefore, even though closely related DAS28 measures for 2 patients might fall either side of an arbitrary threshold they might actually relate to patients who are phenotypically very similar. Presuming that comparisons of groups with the greatest phenotypic difference would be most likely to elicit differences in underlying gene expression profile, the baseline and 3 month DAS28 scores were used to form participants into sub-groups with clear, statistically distinct, differences in the DAS28 score. Overall, none of the comparisons of different levels of disease activity were associated with differences in whole blood gene expression at either baseline or after 3 months of DMARD monotherapy. However, this may have been due to the grouping structures chosen for each comparison. In order to fulfil the inclusion criteria, participants at baseline had to display at least moderate disease activity (DAS44 >2.4). Hence, the comparison of baseline gene expression profiles in relation to disease activity is actually between the lowest quartile of participants, who all had moderate disease activity (median DAS28 3.57), and the highest quartile of participants, who all had high disease activity (median DAS28 6.48). Despite the distribution of DAS28 scores between the two groups being statistically distinct ( $p < 0.0001$ ) the moderate and high disease activity states might still have been too close to each other to exhibit clear differences in gene expression.

After 3 months of DMARD therapy, the remaining participants exhibited a much wider range of DAS28 scores that encompassed the whole range of potential disease activity states (DAS28 range 1.049 – 6.3). Ideally, the 3 month gene expression profiles of participants in DAS28 LDAS and/or clinical remission would have been compared to those of participants who were still in DAS28 high disease activity. However, after 3 months of follow-up the majority of participants had experienced some benefit from their initial treatments and the DAS28 scores were strongly positively skewed towards DAS28 LDAS and remission. In fact, nearly half had already attained DAS28 remission. Thus comparing groups of participants with high disease activity ( $n=4$ ) and participants with DAS28 remission ( $n=35$ ) would have been substantially imbalanced. Therefore, to allow a comparison between 3 month disease activity states participants were separated into comparator groups using the upper and lower quartiles of the 3 month DAS28. This did produce comparator groups that appeared phenotypically distinct. Patients in the lower quartile group all had DAS28 remission (DAS28 median 1.73), whilst patients in the upper quartile group all had moderate-high disease activity (DAS28 median 4.46). Nevertheless, despite clear phenotypic differences there was no evidence of differential gene expression between the comparator groups. Further attempts were made to increase the sensitivity of the gene expression analysis by forming comparator groups comprising participants with even wider differences in clinical phenotype. Several different manipulations of the available DAS28 and MSUS disease activity data were made to allow the formation of several different phenotypic extreme comparator groups. None of these approaches however identified evidence of differential gene expression despite the comparator groups exhibiting clearly different levels of disease activity.

An alternative method of performing comparisons of phenotypic extremes would have been to form a comparator groups comprising patients who had attained the stringent ACR-Boolean definition of remission (186), since this may be a more clinically relevant for the majority of rheumatologists who cannot perform MSUS assessments. However, this comparison was deferred since: 1. the preceding two comparisons based upon DAS28 remission and MSUS LDAS/remission had failed to identify any signal that extremes of disease activity would be associated with gene expression differences and 2. ACR-Boolean Remission did not necessarily equate to absence of disease activity. Twelve participants attained ACR-Boolean Remission and also underwent MSUS disease activity assessment at the 3 months time point. Of these participants, only 3 (25%) exhibited total absence of PD signal; 3 had a PD joint count of 1 and 6 (50%) had evidence of active disease with a PD joint count of 2 or higher (range 2-4) (data not shown).

Clinically, at a cohort level, there was a positive response to initial DMARD monotherapy, best evidenced as a statistically significant fall in the mean DAS28 score between baseline and 3 months. However, not all participants exhibited a significant improvement in DAS28; in 5 participants DAS28 fell by less than 1.2 (the minimum clinically important change) and in 8 participants DAS28 actually increased. It is possible that differences in magnitude, and direction, of treatment response may also have been reflected in out-of step changes in gene expression that skewed the outputs of comparisons involving the whole cohort. A crude, unadjusted, comparison was conducted between baseline and 3 month gene expression profiles using data from all participants and demonstrated that 19 genes underwent changes of expression between baseline and the 3 month time point. It is possible that the differential expression of these genes represents a true dynamic change in response to treatment. However, similar differences were not identified when the baseline and 3 month gene expression patterns were compared in the subgroup of participants who exhibited the greatest phenotypic change over the same time period (i.e. DAS28 remission, MSUS LDAS/remission). Since these participants had experienced the biggest change in disease activity it was presumed that they would be most likely to also demonstrate a dynamic change in gene expression. Further restriction of the comparison to the 63 participants who exhibited a significant improvement in DAS28 between baseline and 3 months (i.e. fall in DAS28 of 1.2 or greater) may reveal additional changes in gene expression profile that are not evident at the cohort level. Given the breadth of the microarray and the size of the cohort it is quite possible that the interplay between the various external factors that might also influence gene expression (e.g. initial DMARD therapy, corticosteroid use, variations in clinical course, intercurrent illness) caused certain genes appearing to appear to undergo expression changes which, whilst temporally associated with the commencement of treatment, were not strictly disease activity related.

Whilst the early treatment regimen was standardised, differences in DMARD escalation threshold meant that treatment was not applied uniformly to all participants. Equally, the quantity and frequency of corticosteroid administered to each participant varied quite considerably depending upon their clinical course and month-by-month disease activity level. Since differences in gene

expression have been reported in relation to methotrexate (374,375), sulphasalazine (357) and corticosteroids (359,361) it is possible that the variable treatment exposure amongst participants contributed to additional variations in gene expression patterns that obscured any expression patterns that were directly related to disease activity. It is possible to correct for some of the treatment-related influence on gene expression by limiting comparisons to participants who solely received methotrexate. However, it would still be difficult to fully correct for the influence of concomitant corticosteroid exposure and to disentangle whether any measured changes in gene expression were related to changes in disease or the treatment exposure.

The principle findings of previously published studies comparing RA disease activity and gene expression can be summarised thus:

1. Treatment with either rituximab (356) or anti-TNF $\alpha$  blocking (376) therapy induces detectable changes in peripheral blood gene expression
2. Characteristic patterns of gene expression from within fibroblast synovial-like cells (377) and synovial tissue (371) correlate with different disease activity states, as measured by either HAQ, DAS28, ESR or CRP
3. Specific PBMC gene expression signatures have been associated with different levels of disease activity based upon either an absolute DAS28 threshold (370) or subjective assessment of need to escalate therapy (378).
4. Non-response to both methotrexate and infliximab has been individually associated with specific PBMC gene expression patterns that are not present in responders (375)

Therefore, the indications are that synovial tissue and peripheral blood gene expression patterns might in some way correlate to RA disease activity and could feasibly serve as an alternative disease activity measure in certain circumstances. However, this tentative consensus is not in keeping with the findings of this research which found that:

1. Whilst there was evidence of change in peripheral blood gene expression profile between baseline and after 3 months of initial DMARD therapy, it could not be reproduced in the specific subsets of participants who experienced the greatest phenotypic improvement
2. Comparisons between upper and lower DAS28 quartile groups at baseline and 3 months did not identify significant evidence of differential gene expression in PBMCs

Allowing for the limitations of the chosen comparator groups these findings do not suggest that it is presently possible to describe a consistent association between peripheral blood gene expression and clinically relevant RA disease activity subtypes. A potential role for gene expression profiling might be to differentiate between patients with active and inactive RA or to confirm response to a particular therapy. The main role for such measures is usually in patients with low levels of inflammatory disease activity where clinical assessments are equivocal and it is difficult to differentiate between those with inactive disease (who do not require DMARD escalation) and those with subclinical active disease (who may require DMARD escalation).

However, if the gene expression measure is no different between patients with the widest phenotypic difference it is unlikely to be sufficiently sensitive or specific to differentiate between those with much subtler differences in disease activity. Junta et al have reported that there was evidence of differential gene expression between relatively small groups of RA patients who fell either side of an arbitrary DAS28 threshold of 5.0. Whilst this is suggestive that higher disease activity is associated with specific patterns of gene expression there is no clear biological reason why the threshold should be a DAS28 of 5.0. Given the size of the study it is possible that a small number of aberrant gene expression data sets could have skewed the overall findings. Galligan et al and van Baarsens et al have also demonstrated that particular gene expression patterns from synovial tissue (van Baarsen (371)) and synovial tissue derived fibroblast-like cells (Galligan et al (377)) do correspond to particular histological and clinical measures of disease activity. However, importantly, van Baarsen et al could not identify corresponding evidence of differential gene expression in simultaneous peripheral blood samples (371). This strongly suggests that synovial tissue, which directly represents the main inflammatory lesion of RA, may be a more appropriate tissue in which to identify evidence of an association between gene expression and disease activity. Whilst peripheral blood is more readily accessible than synovial tissue it may be exposed to too many additional external influences to provide an accurate representation of the nature of gene expression patterns in RA. Equally, RA-related variations in gene expression that occur within the synovium may be either isolated from the systemic circulation or of too small a magnitude to be apparent in whole blood.

In the hunt to identify additional disease activity measures, there remains a tension between identifying a tissue whose gene expression patterns accurately and consistently reflect disease activity and the need for the necessary tissue collection procedures to be acceptable to the patient. Ultimately, it may be that specific patterns of synovial tissue gene expression are found to serve as accurate measures of RA disease activity; however, the clinical relevance of the measure will be significantly limited if patients consider the need for synovial biopsy excessively invasive. Peripheral blood provides an ideal medium for identifying additional disease activity measures since it is readily accessible and favours repeated sampling to track fluctuations in disease activity. However, the results from this research suggest that, despite its accessibility, peripheral blood may not be an appropriate tissue to investigate since, so far, differences in disease activity have not been consistently associated with evidence of differential gene expression. Taken together the inherent limitations of both synovial tissue and peripheral blood gene expression analysis suggest that it might not be possible to develop a reliable and accurate RA disease activity measure based upon gene expression patterns; though they may still continue to contribute important information to the understanding of disease pathogenesis and subtyping.

Unfortunately, due to evident skew in the distribution of DAS28 values at baseline (negative skew towards moderate and high disease activity) and follow-up month 3 (positive skew towards low disease activity and/or remission) it was not possible to create a single comparison that represented the whole range of DAS28-defined disease activity at a single time point. Whilst a clinical trial environment is a very good setting for analysing change in gene expression in

response to a specific intervention (because the influence of potential confounders can be anticipated and minimised) it does inevitably lead to skew in the distribution of disease activity measures at baseline and after commencement of treatment because the overriding treatment aim is to suppress inflammatory disease activity as quickly as possible. An alternative approach would be to either 1. investigate a longitudinal early RA cohort or 2. recruit and assess patients attending routine follow-up clinics in the hope of capturing the whole spectrum of RA disease activity. However, both these approaches were outwith the logistical scope of this research.

### **7.6.3 Relationship between treatment response and baseline gene expression profile**

This research identified 3 genes (TBC1D22B, TMCC2, USP46) that were up-regulated in baseline samples of participants who failed to achieve DAS28 LDAS after 3 months of DMARD monotherapy. However, these genes were not identified until the adjusted p value threshold had been increased to 0.15. Furthermore, differential expression of the same genes was not identified when the participants were grouped according to quartiles of the 3 month DAS28 value. In fact, there was no apparent association between baseline gene expression profile and any particular categorisation of 3 month disease activity. Overall, on the basis of this study it is not possible to conclude that a particular baseline gene expression profile is neither associated with, nor potentially predictive of, disease activity level after 3 months of DMARD monotherapy. The up-regulation of 3 genes at baseline may differentiate between those patients who are likely to require DMARD escalation after 3 months of DMARD monotherapy; however, this would require validation through quantitative PCR analysis of residual samples from this cohort and prospective validation in an independent RA cohort.

Participant's mean DAS28 between baseline and follow-up month 12 was used as a surrogate measure of their disease course during the first year of treatment. The upper and lower quartile groups had evidently experienced quite different inflammatory disease burdens; for the lower quartile group the median DAS28 0-12 months was 1.93 (suggesting persistent LDAS, if not remission), whilst for the upper quartile group the median DAS28 0-12 months was 4.0 (suggesting moderate – high disease activity). Even so, the evident differences in disease course were not associated with differences in baseline gene expression profile. Therefore, based on these results, it is not possible to use baseline whole blood gene expression profile to predict a patient's likely disease course, nor their potential requirement for a specific intensity of DMARD therapy. The EULAR response criteria for RA (297) do provide a categorical measure of treatment response over time and could have been used as an alternative grouping variable for disease course. However, they are based upon the final value of DAS28/DAS44 and the relative change from another point and can only represent the change in disease activity level between those two time points. In RA, where month-by-month disease activity levels can be dynamic and highly variable, the relative change between the two points may not actually reflect the overall disease activity burden in the intervening month. Hence, for this comparison, the mean DAS28 value was chosen as the comparator variable since it does incorporate all of the intervening DAS28 values.

To date, most studies have correlated pre-treatment gene expression profiles to short term clinical outcomes in relation to single drug interventions. Previous studies have demonstrated that specific patterns of pre-treatment gene expression in the whole peripheral blood, PBMC and synovial biopsies predict the likelihood of response to either infliximab (234,236,370,371,379), adalimumab (239,380), rituximab (355,356,381) and anakinra (382). Several studies have also attempted to maximise the practical application of gene profile assessment by identifying either single genes (CD11c, (239)), reduced gene profiles (234,355,356) or mean gene expression scores based upon a small number of candidate genes (355,356). However, the generalizability of these results remains uncertain. The findings of most studies require further validation through replication of the gene signature by qRT-PCR and/or further testing in a larger, prospective cohort.

Two studies have described how changes in gene expression profile after commencement of anti-TNF $\alpha$  blocking therapy may also predict the likelihood of attaining a favourable response (238,383,384). Whilst both studies identified changes in gene expression profile that were associated with subsequent treatment response it is unclear whether this approach will be clinically useful since: i. it requires patients to undergo a two step tissue sampling process to assess the change in gene expression profile from baseline and ii. it still requires patients to be exposed to an agent, which may ultimately still prove ineffective, whilst the gene expression response is being assessed.

Both the studies by Raterman et al and Vosslamber et al have suggested that pre-treatment interferon response gene scores, and an increase in interferon response gene expression after 3 months, are potentially predictive of achieving a good response to rituximab and that dynamic early changes in interferon gene expression may reflect clinical response (355,356). The clinical applicability of these scores is evident, since the mean interferon response gene score is both objective and easily interpreted. However, the findings of the studies by Raterman et al and Vosslamber et al are so far isolated to treatment with rituximab and therefore of restricted clinical value. Interestingly, in this study, pre-treatment mean interferon response gene score and the 3 month IRG T3:T0 ratio were not found to have the same relationships with clinical response to step-up DMARD therapy after either 3, 6 or 12 months. In fact, the value of 3 month IRG T3:T0 ratio for the majority of participants in this research was very close to 1.0, suggesting that there was virtually no change in interferon gene expression following commencement of DMARD monotherapy. Therefore, the results from this study suggest that interferon response gene expression scores, are unlikely to serve as either predictive markers of short-medium term response or additional measures of disease activity in RA patients receiving step-up DMARD therapy.

The majority of the studies that have so far described gene expression profiles that are potentially predictive of treatment response, have been conducted using relatively small patient cohorts (n=4-62) where there is a higher risk that aberrant results may skew the findings. By contrast, the cohort investigated by this study was notably larger (n=79) than that used in most previous

studies. Whilst this increases the risk of multiple testing introducing false positive results it should also attenuate the risk of aberrant results adversely influencing the overall results. However, despite the cohort's size, this study has not identified a consistent relationship between pre-treatment gene expression profile and either 3 month disease activity level or clinical course over 12 months. Interestingly, a similar pattern was observed in comparisons between synovial tissue gene expression and treatment response. The smaller studies reported by Badot et al and van der Pouw Kran et al (n= 18 and 25 respectively) had identified evidence of differential gene expression between responder and non-responder patients (236,380). However, a subsequent larger study by Lindberg et al (n= 62) did not identify any differential gene expression in the baseline profiles, except in a subset of patients with histological evidence of synovial lymphoid aggregates (235).

The disconnect between the findings of small and large gene expression cohort studies was also evident when this study tested whether the mean interferon response gene scores proposed by Raterman et al and Vosslamber et al exhibited the same relationship to short-medium term response in patients receiving step-up DMARD therapy (355,356). Admittedly the parameters of the comparison were different since 1. DAS28 data from several different follow-up time points was used instead of just 6 months; 2. participants in this research had early RA, rather than established RA, and 3. they were treated with step-up DMARD rather than rituximab.

Nevertheless, interferon mediated pathways have recently been implicated in the development of symptomatic disease in ACPA and RhF positive RA patients (240,385,386) so it is reasonable to presume that the importance of interferon response genes will not be solely confined to the likelihood of achieving a response to rituximab. Eitherway, this research was unable to replicate the findings of either study since there was no significant difference in baseline mean interferon gene scores between participants categorised as responders and non-responders after either 3 or 12 months follow-up and there was little evidence of a dynamic change in interferon response gene expression following commencement of DMARD therapy. The influence of intercurrent illness and co-morbidities on peripheral blood gene expression profiles may also ultimately limit the value of the interferon response gene signature in RA, particularly since elevated expression of interferon response genes has been reported in systemic sclerosis (387), juvenile idiopathic arthritis (388), systemic lupus erythematosus, multiple sclerosis and type 1 diabetes mellitus(389).

All previously published studies have described the association between treatment response and pre-treatment gene expression profile in RA patients commencing a single immunomodulatory agent (usually a biologic agent). Thus, the results of individual studies can at present only be interpreted in relation to the associated therapeutic agent. Whilst there appears to be very little overlap between the various gene profiles that have been reported it cannot be presumed that the temporal association with treatment response is confined to either that agent or indeed RA. Prior to routine clinical use, the extended performance characteristics of individual gene profiles will require testing in alternative clinical circumstances (e.g in relation to alternative diagnoses and/or drug treatments) to determine their specificity and sensitivity to different clinical scenarios, and therapeutic agents, rather than the restrictive inclusion criteria of a clinical study. Feasibly, it may

be possible to identify pre-treatment gene expression profiles that reliably predict response to each individual immunomodulatory agent. Provided there is minimal overlap of the constituent genes, the predictive profiles for several different agents could be combined in a single microarray that could be used to estimate an individual's likelihood of responding to a range of potential therapeutic options; therefore, facilitating the clinician choosing the agent with the greatest likelihood of success. Microarrays that comprise separate predictive gene profiles for several immunomodulatory agents are more likely to be clinically relevant since they will maximise the number of occasions that the microarray is indicated, rather than being confined to single occasions when only one agent is being considered.

Since the majority of PAXgene RNA samples were collected prior to the commencement of DMARD treatment, variations in the choice of first DMARD and use of corticosteroid could not have influenced the baseline gene expression profiles. The 3 month assessment point allowed investigation of the ability of pre-treatment gene expression profiles to predict response to initial DMARD monotherapy (predominantly methotrexate). For clinical and societal reasons, a number of participants initially commenced sulphasalazine instead of methotrexate. Hence, a further, truer pharmacogenomic analysis comprising only the pre-treatment gene expression profiles of the subset of participants who commenced methotrexate monotherapy (n=74 of 79) is also planned. Comparisons of 12 month disease activity level and mean DAS28 to pre-treatment gene expression profiles can not be related to a single agent; instead the gene expression profile must be considered in the context of the overall treatment strategy which could potentially include several different non-biologic and biologic DMARD agents. A 12 month assessment point was chosen since it fitted conveniently within the timescale available to conduct the analysis and the longer period increased the likelihood of changes of the DAS28 being able to represent the overall effectiveness of the strategy. However, the complexity of the escalation strategy, and the length of the follow-up period also increased the likelihood of variability within the DMARD therapy received by each patient. Variations in application of the DMARD protocol caused by unexpected adverse events, intercurrent illness and/or patient choice, could have produced variations in disease activity and clinical response in some patients that lead to their pre-treatment gene expression profile data being classified within an alternative comparator group, thereby, potentially skewing the whole group's gene expression data set.

The maximum follow-up period for the transcriptomic analysis (12 months) was longer than that generally reported by previous studies (usually 4-6 months). In previous studies, it is possible that relatively short follow-up periods may have prevented the full extent of the treatment response becoming apparent (particularly in relation to rituximab), and may have lead to individual patients treatment response being misclassified. The longer follow-up period of this study may have increased the sensitivity of the clinical assessment to detect change in disease activity following commencement of treatment. However, the differences in follow-up period between this research and previous studies do address slightly different clinical scenarios. Being able to accurately predict short-term treatment failure of a single agent (especially biologic agents) may steer a treatment decision towards alternative agents that are more likely to produce a beneficial response. Alternatively, using baseline factors to predict medium term response to a treatment

strategy may have important prognostic properties and could lead to patients being stratified towards different intensities of treatment based upon their prognostic assessment.

**8. Disease Activity Assessment  
Using a Multi-Biomarker  
Disease Activity Test**

## 8.1 Introduction

This chapter will discuss how a cytokine-based measure of disease activity relates to corresponding MSUS and DAS28 disease activity assessments. The laboratory analysis was conducted in collaboration with Crescendo Biosciences, San Francisco, California ([www.crescendobio.com](http://www.crescendobio.com)).

### 8.1.1 The Multi-Biomarker Disease Activity Test

The Multi-Biomarker Disease Activity Test (MBDA) was devised on the premise that the level of inflammatory biomarkers within the blood would be directly related to the level of inflammatory disease present (383). It was hypothesised that it would be possible to create a single objective score of RA disease activity, based on the expression levels of inflammatory biomarkers, that was proportionate to RA disease activity and would allow longitudinal monitoring of disease activity. Serum samples for 649 RA patients (early and established disease) were collected from the collections of 4 existing inception cohort RA studies. Samples were screened for the presence of 130 candidate biomarkers that had previously been associated with either RA disease activity and/or pathogenesis. Univariate and multivariate analyses were conducted for each study individually to determine which markers most closely related to RA disease activity. Subsequently, additional multivariate algorithm development techniques were applied to the biomarker ranks to generate summative formulae that predicted each of the individual components of the DAS28, except CRP. Eventually, a 12-biomarker model was proposed based upon the following biomarkers:

Adhesion molecules	Vascular cell adhesion molecule-1 (VCAM-1_
Growth factors	Epidermal growth factor (EGF)
	Vascular endothelial growth factor A (VEGF-A)
Cytokine-related	Interleukin-6
	Tumour necrosis factor receptor, type 1 (TNF-RI)
Matrix metalloproteinases	Matrix metalloproteinases-1 (MMP-1)
	Matrix metalloproteinase-3 (MMP-3)
Skeletal proteins	YKL-40
Hormones	Leptin
	Resistin
Acute-phase proteins	Serum amyloid (SAA)
	C-reactive protein (CRP)

The expression levels of the 11 candidate biomarkers (excepting CRP) were combined in summative formulae that estimated the value of the corresponding 28 tender joint count, 28 swollen count and patient 10cm global VAS (390):

$$\text{predicted 28 tender joint count} = -26.72 + 3.243 * [\text{YKL-40}]^{1/10} - 11.97 * [\text{EGF}]^{1/10} + 15.72 * [\text{IL6}]^{1/10} + 0.4594 * [\text{Leptin}]^{1/10} + 3.881 * [\text{SAA}]^{1/10} + 0.7388 * [\text{TNF-RI}]^{1/10} - 0.2557 * [\text{VCAM1}]^{1/10} + 0.7003 * [\text{VEGF-A}]^{1/10}$$

$$\text{predicted 28 swollen joint count} = -26.63 + 3.232 * [\text{YKL-40}]^{1/10} - 11.93 * [\text{EGF}]^{1/10} + 15.67 * [\text{IL-6}]^{1/10} + 0.4578 * [\text{Leptin}]^{1/10} + 3.868 * [\text{SAA}]^{1/10} + 0.7363 * [\text{TNF-RI}]^{1/10} - 0.2548 * [\text{VCAM-1}]^{1/10} + 0.6979 * [\text{VEGF-A}]^{1/10}$$

$$\text{predicted global 10cm VAS} = -13.489 + 5.474 \cdot [\text{IL-6}]^{1/10} + 0.486 \cdot [\text{SAA}]^{1/10} + 2.246 \cdot [\text{MMP-1}]^{1/10} + 1.684 \cdot [\text{Leptin}]^{1/10} + 4.14 \cdot [\text{TNF-RI}]^{1/10} + 2.292 \cdot [\text{VEGF-A}]^{1/10} - 1.898 \cdot [\text{EGF}]^{1/10} + 0.028 \cdot [\text{MMP-3}]^{1/10} - 2.892 \cdot [\text{VCAM-1}]^{1/10} - 0.506 \cdot [\text{Resistin}]^{1/10}$$

Thereafter, the outputs of each of these formulae were incorporated with the CRP into a derivation of the established DAS28 formula to provide the final MBDA score:

$$\text{MBDA} = \text{round} [\max (\min ((0.56 \sqrt{\text{pTJC28}} + 28 \sqrt{\text{pSJC28}} + 14 \text{pPGAz} + 36 \ln(\text{CRP}+1) + 0.96) \times 10.53 + 1, 100), 1)]$$

The MBDA calculation outputs a single numerical value in the range 0-100 and MBDA thresholds have been proposed that correspond to existing DAS28-CRP disease activity thresholds (Table 52). At present there is no accepted definition of minimal clinically important difference.

Disease Activity Category	DAS28-CRP Definition	MBDA Definition
Remission	<2.3	≤25
Low	≥2.3 and ≤2.7	>25 and ≤29
Moderate	>2.7 and ≤4.1	>29 and ≤44
High	>4.1	>44

**Table 52** - Disease activity category definitions based upon DAS-CRP and MBDA (after Centola et al 2013 (383))

#### *Validation of the Multi-Biomarker Disease Activity Test*

During its initial derivation, the performance of the MBDA was tested in an independent study of 24 RA patients who were receiving treatment with methotrexate and infliximab and underwent serial assessment of disease activity using DAS28, MSUS (MCP joints only) and the MBDA (391). Overall, the MBDA performed favourably compared to other measures of disease activity (383). There was significant positive correlation between the MBDA and corresponding DAS28-CRP (Spearman's  $r=0.82$ ,  $p<0.001$ ), MSUS evidence of synovial hypertrophy (Spearman's  $r = 0.46$ ,  $p<0.001$ ), MSUS PD area (Spearman's  $r = 0.47$ ,  $p<0.001$ ) and change in DAS28 (Spearman's  $r = 0.62$ ,  $p<0.001$ ). Furthermore, baseline MBDA score showed a stronger correlation (Spearman's  $r = 0.52$ ) with 12 month change in total Sharp score than DAS28-CRP (Spearman's  $r = 0.43$ ,  $p=0.006$ ).

The performance of the MBDA has been tested in an independent validation exercise conducted using samples from an additional 230 RA patients who were participating in 3 existing RA prospective cohort studies (390). Correlation analyses demonstrated that the MBDA score correlated positively with DAS28 in both sero-positive and sero-negative RA patients (Pearson's  $r = 0.56$  and  $0.43$  respectively,  $p<0.001$  for both). There was fair-moderate agreement between the DAS28-CRP and MBDA categorization of disease activity state for both sero-positive (kappa  $0.44$ ,  $p<0.001$ ) and sero-negative (kappa  $0.33$ ,  $p<0.001$ ) patients. In sero-positive and sero-negative patients, the MBDA correlated significantly with other composite disease activity measures; however, the strength of the correlation appeared consistently stronger in sero-positive patients.

The MBDA score's responsiveness to changes in disease activity following treatment has been tested in an independent analysis of 45 RA patients who were all treated with methotrexate and

infliximab. A positive response was defined as a fall in DAS28-CRP of 1.2 or more. In treatment responders, the MBDA score demonstrated significant improvements between baseline and week 12 of treatment, with the greatest rate of change occurring in the first 2 weeks. In non-responders, the changes in MBDA score were much less marked. The change in MBDA between baseline and week 2, positively correlated with the corresponding change in DAS28-CRP (Spearman's  $p = 0.51$ ,  $p < 0.001$ ) and showed a fair ability to discriminate between DAS28-CRP responders and non-responders after 2 weeks (AUROC = 0.72,  $p = 0.02$ ) and 12 weeks (AUROC = 0.77,  $p = 0.002$ ) of follow-up.

Two previously discussed DMARD strategy studies have also recently described the correlation between the MBDA score and DAS28 defined disease activity in patients with RA. In both cases, MBDA was calculated using stored serum samples and was compared to existing clinical outcome datasets. Bakker et al compared the relationship between MBDA and DAS28-CRP in 72 patients recruited to the CAMERA study (392). Seventy-two sample sets were available at baseline; however, only 48 samples were available after 6 months follow-up. The MBDA score correlated strongly with the DAS28-CRP (Pearson's  $r = 0.72$ ,  $p < 0.001$ ) and could distinguish between remission/LDAS and moderate/high disease activity (AUROC 0.86,  $p < 0.001$ ). However, there was only fair agreement between the MBDA and DAS28-CRP disease classification for each disease activity subtype (kappa 0.34, 95%CI 0.19-0.49) and MBDA was not shown to be a strong predictor of subsequent radiographic progression over the next 2 years. In a similarly designed study, Hirata et al described the correlation between DAS28-ESR and MBDA in a subset of 125 RA patients recruited to the BeST study (393). Follow-up data after one year was available in 54 patients (43%) who had been selected from all of the original studies treatment strategy arms. Once again, the MBDA score correlated positively and strongly with DAS28-ESR (Spearman's  $r = 0.66$ ,  $p < 0.0001$ ), SDAI (Spearman's  $r = 0.67$ ,  $p < 0.0001$ ) and CDAI (Spearman's  $r = 0.56$ ,  $p < 0.0001$ ). Further, the change in MBDA after 1 year's follow-up correlated positively with the change in DAS28-ESR (Spearman's  $r = 0.55$ ,  $p < 0.0001$ ) and SDAI (Spearman's  $r = 0.35$ ,  $p = 0.016$ ), but not CDAI (Spearman's  $r = 0.18$ ,  $p = 0.23$ ), over the same time period.

As previously discussed, DAS28 assessment of disease activity is insensitive to some subclinical synovitis and DAS28 remission does not necessarily predict non-progression of radiographic damage (121,200). Further, whilst the DAS28-CRP and MBDA have demonstrated positive correlations at a single time point, and in relation to treatment changes, it is possible that the MBDA might provide a more accurate marker of ongoing subclinical synovitis. Therefore, it is also possible that MBDA remission might prove a better marker of true remission, and radiographic non-progression, than DAS28 remission. To an extent this has been borne out in an observational study of 163 patients with established RA receiving standard DMARD therapy (394). Patients who were in MBDA remission at baseline exhibited lower rates of radiographic progression compared to patients who were not in MBDA remission. Conversely, patients with high MBDA disease activity had the highest rate of radiographic progression. Furthermore, rates of radiographic progression were consistently higher in patients who achieved alternative definitions of remission (e.g. DAS28 or ACR-EULAR Boolean remission) suggesting that MBDA-remission is a stronger predictor of radiographic remission than current composite definitions.

All of the previous studies have suggested that the MBDA test score correlates reasonably well with the DAS28-CRP as either a static measure of disease activity or a dynamic measure over time. The implicit meaning being that if the scores are closely correlated they must also be measuring the same factor. Indeed, several studies have also reported favourable results from receiver operating curve analyses (range AUROC = 0.76 – 0.89) (383,395,396) in relation to DAS28-CRP thresholds. Interestingly though, no studies have yet reported the sensitivity and specificity of the MBDA test in relation to DAS28 or any other disease activity measure.

The potential impact on clinical practice of the additional disease activity information provided by the MBDA test has been tested in a theoretical and real world setting. Eighty-one rheumatologists were asked to comment on the management of 3 simulated RA patients; 37 (46%) of the rheumatologists were randomized to also receive the corresponding MBDA score (397). The quality of the clinician's performance in different domains of treatment decision making was quantified using a standardized clinometric tool, the Clinical Performance and Value vignettes. Overall, clinicians with knowledge of the MBDA score improved their 'quality scores' by a significantly greater amount than those with no knowledge of the MBDA. However, abstract theoretical decisions can often differ significantly from real-world clinical practice. Six rheumatologists were surveyed regarding their opinion of RA disease activity and need for treatment escalation before and after being made aware of the MBDA result (398). Knowledge of the MBDA result was associated with changes to the treatment plan in 38 of 101 patient's cases. In most cases, changes lead to therapy escalation with either addition of new DMARDs, dose changes and/or changes to the route of administration.

### **8.1.2 Objective**

To determine the performance characteristics of the MBDA test in relation to corresponding DAS28 and, where available, MSUS assessments of global disease activity in RA patients recruited to the TaSER study

## **8.2 Methods**

*Study Cohort:* All 79 research participants who had contributed serum samples to the TaSER study tissue bank were included in the MBDA analysis. This cohort was identical to the cohort included in the previously described gene expression analysis (Chapter 7). Paired clinical and MBDA data were available for 79 participants at baseline, 76 participants at follow-up month 3 and 71 participants at follow-up month 18.

*Disease Activity Assessments:* All clinical and MSUS assessments of disease activity, relating to the baseline, 3 month and 18 month time points were collated from the TaSER study data set. The DAS28-ESR was chosen as the main measure of clinical disease activity, since it had been the main factor about which DMARD escalation decisions were based during the study period. Whilst the MBDA had mostly been tested in relation to the DAS28-CRP, a nested analysis within the BeST study had shown good, positive correlation ( $r=0.66$ ) between the DAS28-ESR and the

MBDA test result (393). Table 53 summarises the disease activity thresholds for the MBDA, DAS28 and MSUS assessments:

	<b>MBDA</b>	<b>DAS28</b>	<b>MSUS</b>
<b>High</b>	>44	>5.1	grade 1 or higher PD signal in 2 or more joints
<b>Moderate</b>	>29 and ≤ 44	>3.2 and ≤ 5.1	
<b>Low</b>	>25 and ≤ 29	>2.6 and ≤ 3.2	PD signal in 1 joint
<b>Remission</b>	≤25	≤2.6	No PD signal in any joint

**Table 53** – Comparison of thresholds for MBDA, DAS28 and MSUS definitions of RA disease activity

*MBDA Analysis:* The laboratory analysis and calculation of MBDA test scores was kindly performed by collaborating colleagues at Crescendo Biosciences. For each participant, 2 of the stored 0.5ml SST serum samples were shipped frozen to the Crescendo Biosciences laboratory. All samples were defrosted simultaneously and analysed en batch according to standard laboratory procedures using the Meso Scale Discovery Multi Array multiplex kits manufactured by Crescendo Bioscience. Array analysis initially returns the concentration of each of the 12 protein markers which are then used to calculate the final MBDA test score using the formulae described in Section 8.1.1.

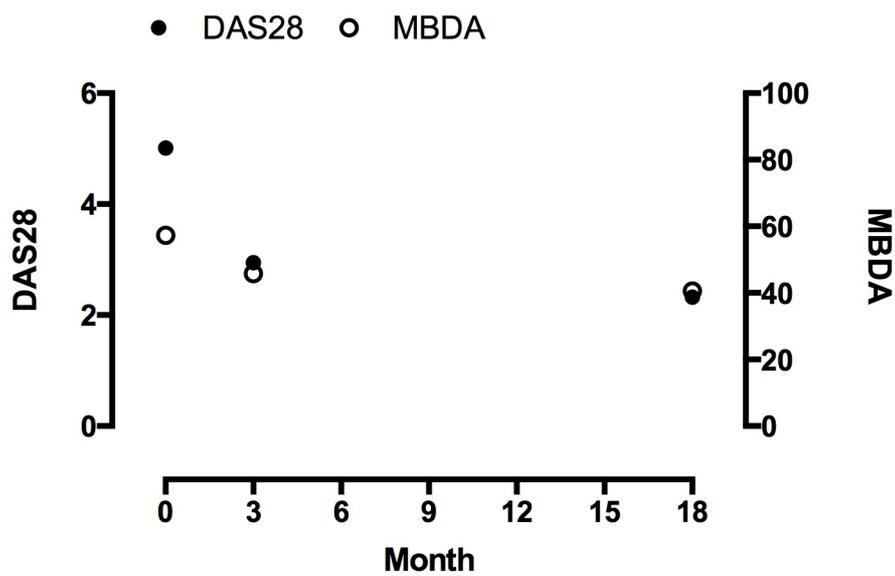
*Statistical Analysis:* Following laboratory analysis Crescendo Biosciences provided the final MBDA score, and the individual concentrations of each component marker, for analysis. All statistical analyses were conducted in Graphpad Prism Version 6 ([www.graphpad.com](http://www.graphpad.com)). The distribution of the difference in MBDA values between baseline and 3 months and baseline and 18 months did follow a Gaussian distribution, therefore paired t tests were used to compare the difference in MBDA between two time points. At each time point the MBDA scores did not demonstrate a Gaussian distribution, therefore correlations between the MBDA and DAS28 were calculated using Spearman's rank correlation. The ability of MBDA to recognise comparable disease activity states to both the DAS28-ESR and MSUS assessments was calculated as the percentage agreement at each time point and over the whole time course. The potential predictive ability of MBDA was estimated by calculating the correlation between baseline MBDA, to mean DAS28 over 12 months. Participants were stratified by randomisation group to determine whether intensity of therapy impacted on the overall change in MBDA score after 18 months.

### 8.3 Results

Results are presented showing the correlation between DAS28-ESR and MBDA at a single time point and also the correlation between change in DAS28-ESR and MBDA over time. The demographic and disease characteristics of the cohort has been previously described in Section 7.1

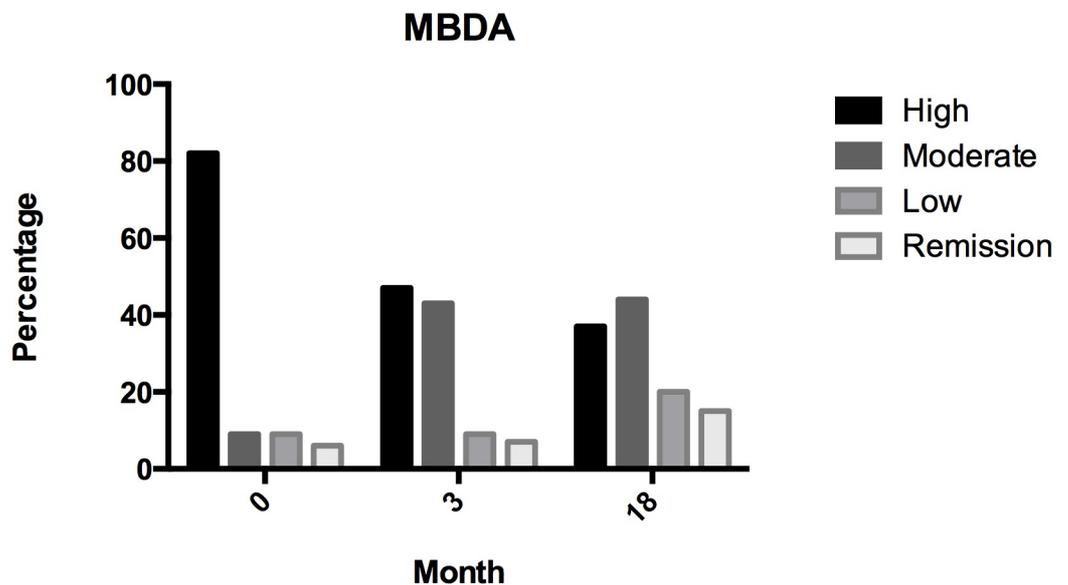
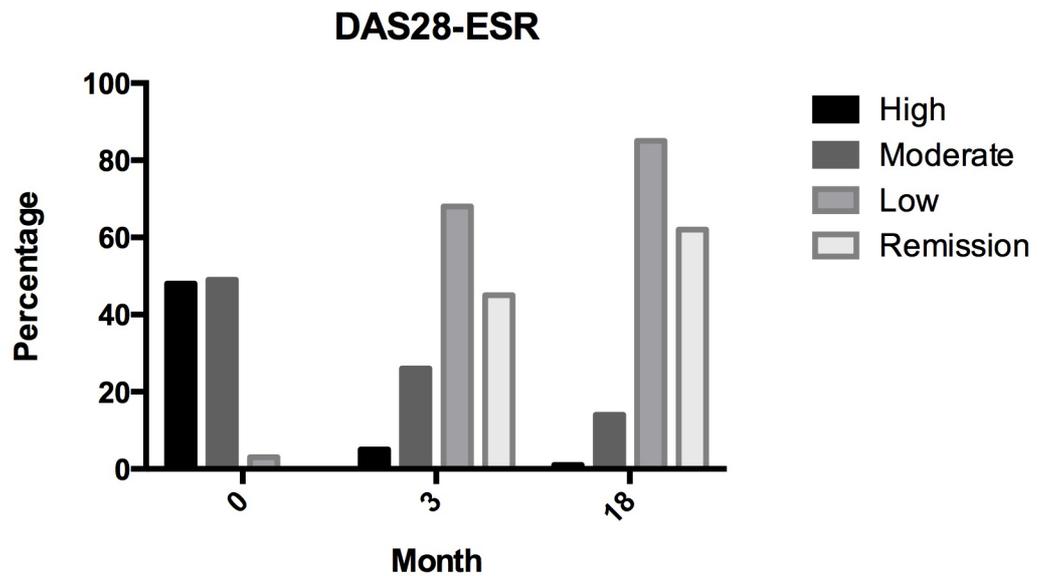
### 8.3.1 Change in DAS28-ESR and MBDA Score

Both DAS28-ESR and MBDA score showed significant improvement over the follow-up period (Graph 53). At baseline mean DAS28 was 5.01 (SD 1.14), after 3 months mean DAS28 was 2.94 (SD 1.21) and after 18 months mean DAS28 was 2.32 (SD 1.03). At baseline, mean MBDA score was 57 (SD 17), after 3 months it was 46 (SD 15) and after 18 months it was 41 (SD 15). Statistical comparisons showed highly significant changes in the values of DAS28 and MBDA between each time point ( $p < 0.001$  for each comparison). The highest rate of change in DAS28 and MBDA was observed between baseline and 3 months.



**Graph 53** – Mean DAS28 and MBDA values at baseline, 3 months and 18 months

For both assessment methods, the proportion of participants classified as high disease activity fell between baseline and 18 months, whilst the proportion of participants classified as low disease activity increased. At each time point, there were striking differences evident in the proportion of participants classified as each disease activity state by DAS28-ESR and MBDA (Graphs 54A and 54B). At follow-up months 3 and 18, the DAS28-ESR classified higher proportions of participants as either LDAS and/or remission whilst the MBDA test was more likely to classify participants as having high or moderate disease activity.



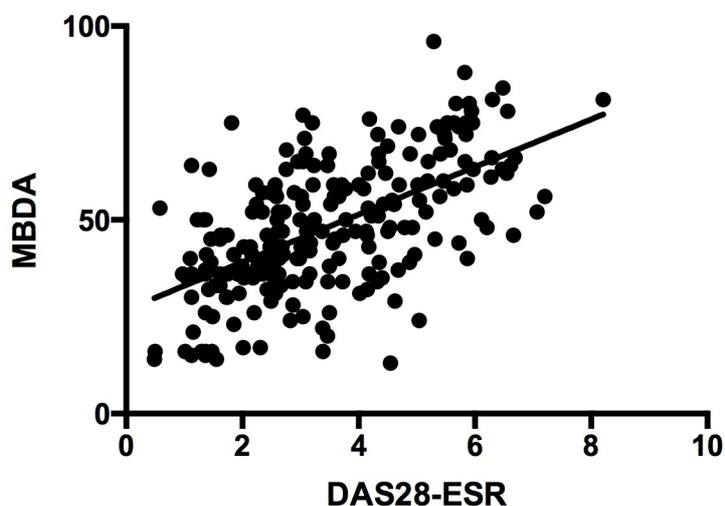
**Graphs 54** – Distribution of disease activity state classification by DAS28-ESR (54A) and MBDA (54B) at baseline, 3 months and 18 months  
**NB** – Low disease activity bar includes remission subgroup

### 8.3.2 Relationship of DAS28-ESR to MBDA Score

Spearman's rank correlation analyses demonstrated that there was a positive correlation between the DAS28-ESR and MBDA at each time point with the strongest correlation evident between the baseline datasets. At baseline, a moderate positive correlation was observed ( $r_s = 0.51$ ,  $p < 0.0001$ ). At 3 and 18 months, low-moderate positive correlation were observed (3 months:  $r_s = 0.37$ ,  $p = 0.001$ . 18 months:  $r_s = 0.48$ ,  $p < 0.0001$ ).

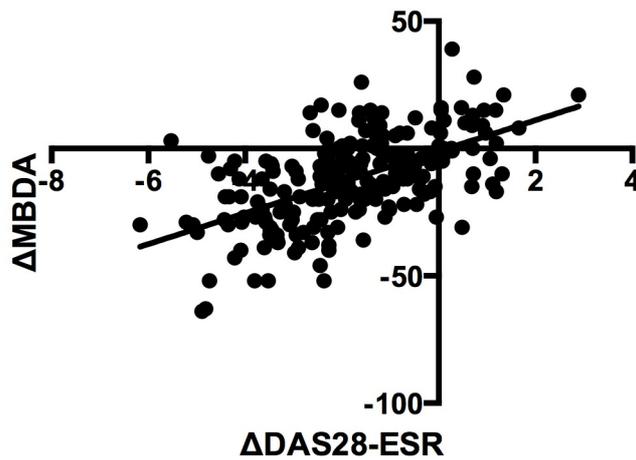
In order to allow a comparison that covered the whole range of DAS28-ESR, various combinations of the baseline, month 3 and month 18 datasets were tested until a combination that most closely resembled a Gaussian distribution of the DAS28-ESR scores was identified. All combinations displayed a degree of positive skew and none matched the typical Gaussian distribution. A combination comprising all available paired sets of DAS28-ESR and MBDA

covered the widest range of DAS28-ESR (0.49 – 8.21) with an acceptable measure of skew (0.39) and the lowest measure of kurtosis (-0.64). The subsequent Spearman's rank correlation analysis between all paired sets of DAS28-ESR and MBDA demonstrated moderate positive correlation ( $r_s = 0.58$ ,  $p < 0.0001$ ) (Graph 55)



**Graph 55** – Relationship between DAS28-ESR and MBDA – all sets of paired DAS28-ESR and MBDA values pooled from each time point

The correlation of the change in DAS28-ESR to change in MBDA was calculated for the time periods between baseline and 3 months, baseline and 18 months and 3 months. Between baseline and 3 months, the mean fall in DAS28-ESR was 2.05 (SD 1.38), the mean fall in MBDA was 12 (SD 16.0) and a moderate positive correlation ( $r_s = 0.48$ ,  $p < 0.0001$ ) was evident. Between baseline and 18 months, the mean fall in DAS28 was 2.68 (SD 1.36), the mean fall in MBDA was 18 (SD 17.6) and a moderate positive correlation was evident ( $r_s = 0.54$ ,  $p < 0.0001$ ). Between, follow-up months 3 and 18 a weaker positive correlation was observed, the mean fall in DAS28 was 0.65 (SD 1.22), the mean fall in MBDA was 5.7 (SD 15.9) and a low-moderate correlation was observed ( $r_s = 0.39$ ,  $p = 0.0007$ ). However, it is evident in all comparisons that the distribution of the data points was not entirely linear; a notable number of outliers corresponded to participants where there was disparity in the degree of change represented by the DAS28-ESR and MBDA (i.e. outlying participants exhibited a marked change in DAS28-ESR but a much lesser change in MBDA). Pooling all of the paired DAS28-ESR and MBDA results together suggested a moderate-strong positive correlation between the change in both measures over the whole follow-up period ( $r_s = 0.56$ ,  $p < 0.0001$ ) (Graph 56). Importantly, a linear regression analysis suggested that the line of best fit passed through the origin.

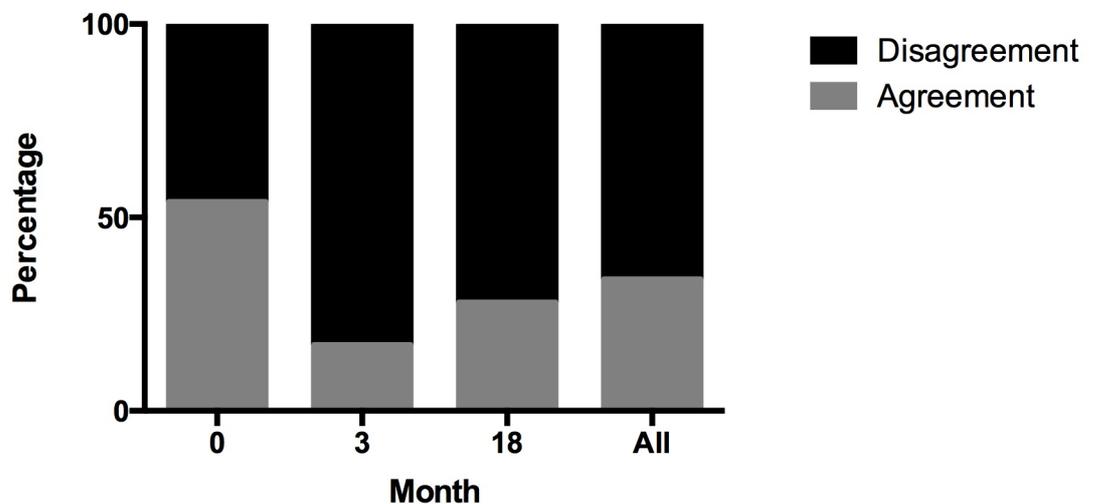


**Graph 56** – Relationship between change in DAS28 and corresponding change in MBDA. The presented results are pooled for the following time periods: baseline-3 months, baseline-18 months and 3-18 months

### 8.3.3 Percentage agreement between DAS28-ESR and MBDA

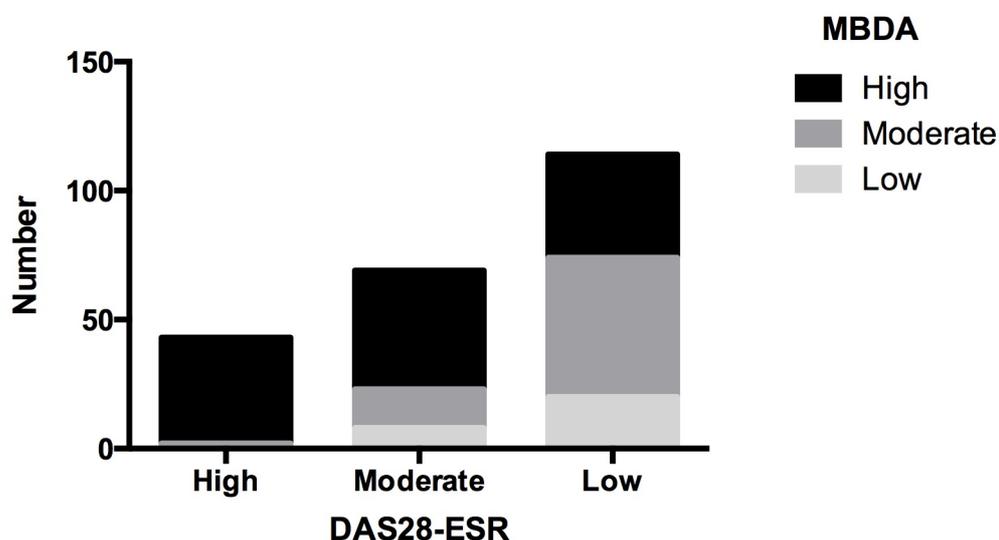
There was significant disparity evident in how each measure was categorising disease activity state for certain participants (Graphs 54A and 54B). Therefore, an analysis was conducted to determine how often the measures agreed on the disease activity state at each time point and over the whole time course. Agreement was defined as both disease activity measures categorising a participant at a comparable level of disease activity.

At every time point, there was significant disparity evident when considering exact agreement between DAS28-ESR and MBDA categorisation of disease activity (e.g. both measures identified high disease activity). Furthermore, the degree of disparity evident increased over the follow-up period (Graph 57). At baseline, there was exact agreement in 43 of 79 assessments (54%), after 3 months there was absolute agreement in 13 of 76 assessments (17%) and after 18 months there was absolute agreement in 20 of 71 assessments (28%). Overall, there was exact agreement in 76 of 226 paired assessments (34%)



**Graph 57** – Percentage exact agreement between DAS28-ESR and MBDA at each time point and pooled for all available results

The majority of instances of disagreement occurred because MBDA identified a higher level of disease activity than initially suggested by the DAS28 (i.e. MBDA was more sensitive than DAS28-ESR). At baseline, there were 36 instances of disagreement; 29 (81%) occurred because MBDA identified a higher disease activity state than DAS28-ESR. At month 3, there were 63 instances of disagreement, in 60 (95%) of these instances the MBDA identified a higher disease activity state than the DAS28. At month 18, there were 51 instances of disagreement, all of which corresponded to the MBDA identifying a higher disease activity state than the DAS28-ESR. Overall, out of 150 instances of disagreement, 140 instances (93%) corresponded to the MBDA test identifying a higher disease activity state than the corresponding DAS28-ESR. Graph 58 summarises the distribution of MBDA disease activity assessment findings when data were grouped based upon the DAS28-ESR disease activity assessment. Similar graphs were produced when the analysis was restricted to each time point (not shown). Of 114 instances of DAS28-ESR low disease activity, 94 (82%) were associated with a higher MBDA disease activity assessment (MBDA moderate = 54, MBDA high = 40) and could potentially have supported DMARD escalation. Of 69 instances of DAS28-ESR moderate disease activity, 46 (67%) were associated with MBDA high disease activity, but would not have altered treatment decisions, and 8 (12%) were associated with MBDA low disease activity. The highest degree of agreement was evident during instances of DAS28-ESR high disease activity since, in all instances, knowledge of MBDA score would not have altered treatment decisions. All 43 instances of DAS28 high disease activity corresponded to either high (41) or moderate MBDA disease activity (2) (i.e. above the treatment escalation threshold).



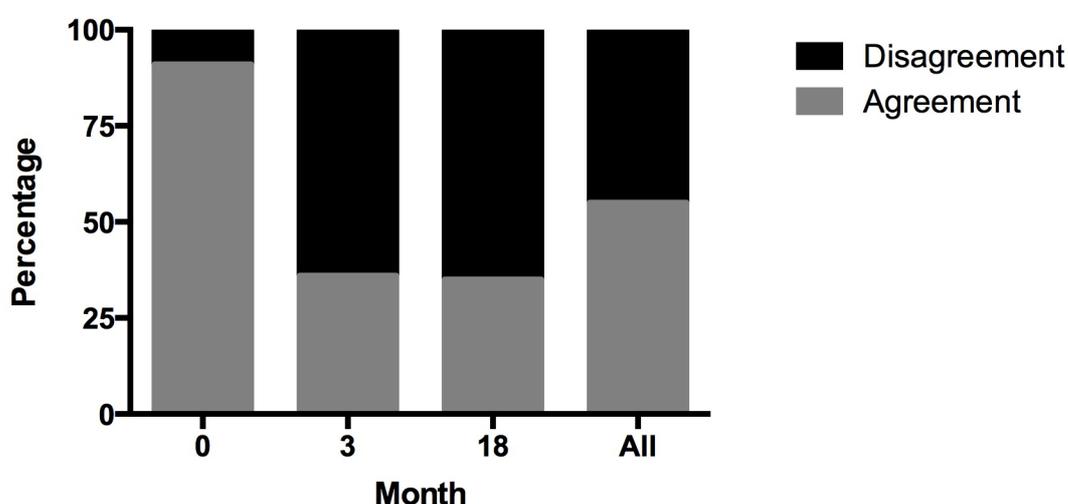
**Graph 58** – Distribution of MBDA disease activity findings when participants were grouped by DAS28-ESR disease activity state. (Results pooled for all time points)

The degree of agreement between DAS28-ESR and MBDA increased when each measure was converted into dichotomous definitions by pooling the findings for closely allied disease activity states (i.e high and moderate disease activity vs low disease activity) (Graph 59). This analysis was based upon the common presumption that moderate disease activity or higher is a suitable threshold at which to consider further DMARD escalation (i.e. DAS28 >3.2). For clarity, comparable DAS28-ESR and MBDA disease activity definitions are shown in Table 54. Using

dichotomous definitions of disease activity states, 72 of 79 assessments (91%) agreed at baseline, 27 of 76 assessments (36%) agreed at 3 months and 25 of 71 assessments (35%) agreed at 18 months. Overall, there was agreement in dichotomous definitions of disease activity states in 124 of 226 assessments (55%). Using DAS28-ESR as reference, the calculated sensitivity of MBDA is 93%, the specificity is 18%, the positive predictive value is 53% and the negative predictive value is 71%.

	DAS28-ESR	MBDA
High-Moderate	> 3.2	> 29
Low-Remission	≤ 3.2	≤ 29

**Table 54** – Comparable dichotomous disease activity states as defined by DAS28-ESR and MBDA



**Graph 59** – Percentage agreement of dichotomous disease activity assessments by DAS28-ESR and MBDA at each time point.

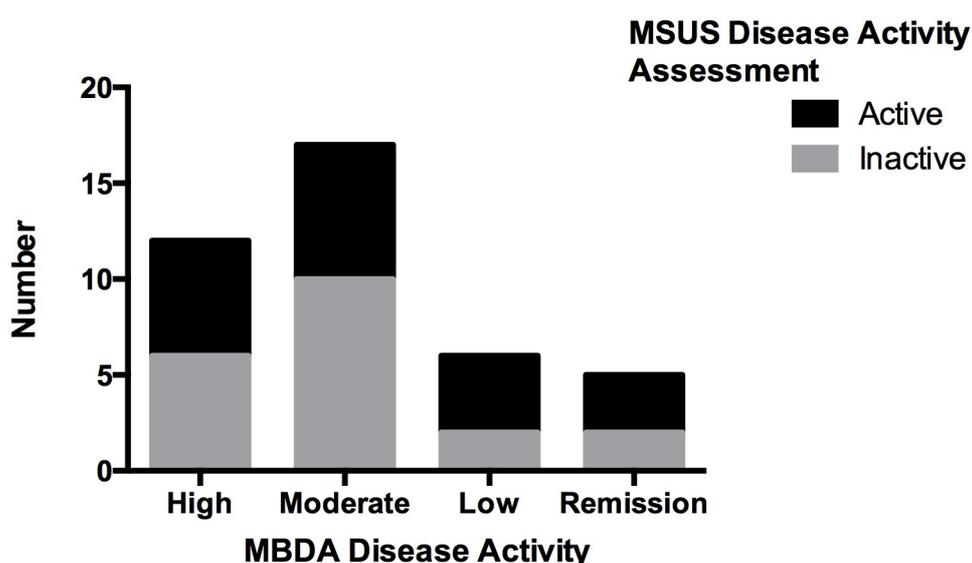
An important aspect of any RA disease activity measure is the ability of the measure to define and detect remission, since this is increasingly the target at which treatment strategies are aimed. At baseline, no participants qualified as DAS28-ESR remission; however, interestingly, 5 participants were classified by the MBDA as remission (DAS28-ESR range = 2.82 – 4.55). At 3 months, 34 participants qualified as DAS28-ESR remission, though only 4 of these participants (12%) also qualified as MBDA remission. Similarly, at 18 months, 44 participants were classified as DAS28-ESR remission, however, only 10 of these (23%) were also classified as MBDA remission. Overall, between months 3 and 18, 78 participants were classified as DAS28-ESR remission, of these 14 (18%) were classified as MBDA remission, 3 (4%) were classified as MBDA low disease activity, 39 (50%) were classified as moderate disease activity and 22 (28%) were classified as high disease activity.

### 8.3.4 Percentage agreement between DAS28-ESR and MSUS assessment

The availability of some concurrent MSUS disease activity information provided an opportunity to test whether the suggestion of on-going inflammatory disease activity by the MBDA was also

reflected in ultrasonographic evidence of active synovitis. This analysis was restricted to the month 3 and month 18 time points. Altogether, there were 35 sets of corresponding DAS28-ESR, MBDA and MSUS data available. Pooled together, the range of DAS28-ESR values were between 0.50 and 4.40. The dataset was skewed towards LDAS and remission; there had been 22 (63%) instances of DAS28-ESR remission, 9 instances of DAS28-ESR LDAS (26%) and 4 (11%) instances of DAS28-ESR moderate disease activity. As before, the MSUS definition of active disease was evidence of grade 1 (or higher) PD signal in at least 2 of the examined joints.

Graph 60 depicts the distribution of MSUS findings when data were stratified by the MBDA disease activity definition. Summaries of the MSUS findings for each MBDA group are shown in Table 55.



**Graph 60** – Musculoskeletal ultrasound findings when participants are stratified by MBDA disease activity definition.

MBDA Disease Activity	High	Moderate	Low
<b>Number</b>	12	17	6
<b>MBDA</b>	54 (IQ 47 – 67)	35 (IQ 35 – 38)	17 (IQ 16.25 – 21.5)
<b>DAS28-ESR</b>	2.9 (IQ 2.6 – 3.1)	2.1 (IQ 1.7 – 2.5)	1.9 (IQ 1.5 – 2.2)
<b>PD Score</b>	3 (IQ 0.75 – 6.25)	1 (IQ 1 – 5)	2.5 (IQ 0.5 – 3)
<b>PD Joint Count</b>	1.5 (IQ 0.75 – 4)	1 (IQ 1 – 3)	2 (IQ 0.5 – 2)
<b>Synovial Hypertrophy Score</b>	4.5 (IQ 2.75 – 7)	3 (IQ 3 – 5)	4 (IQ 2.5 – 4.75)
<b>Synovial Hypertrophy Count</b>	2.5 (IQ 1.75 – 4)	2 (IQ 2 – 3)	2.5 (IQ 1.25 – 3)
<b>Active MSUS Disease – n (%)</b>	5 (42%)	7 (41%)	4 (67%)
<b>PD Joint Count ≥ 1 – n (%)</b>	9 (75%)	14 (82%)	4 (67%)

**Table 55** – Summary of MSUS disease activity findings by MBDA disease activity group. Unless stated, values represent median (IQR)

The high and moderate MBDA groups were combined to differentiate between high-moderate MBDA disease activity and low disease activity in a dichotomous manner. There were 29 instances of MBDA high-moderate disease activity, 13 (45%) of these instances also coincided with active disease during MSUS assessment. There were 6 instances of MBDA low disease activity, 4 (67%) of which corresponded to active MSUS disease. Overall, there was complete agreement between the dichotomous MBDA assessment and MSUS assessment on 15 (43%) occasions. Using the MSUS assessment as reference, the calculated sensitivity of the MBDA as a dichotomous measure was 76%, the specificity was 11%, the positive predictive value was 45% and the negative predictive value was 33%.

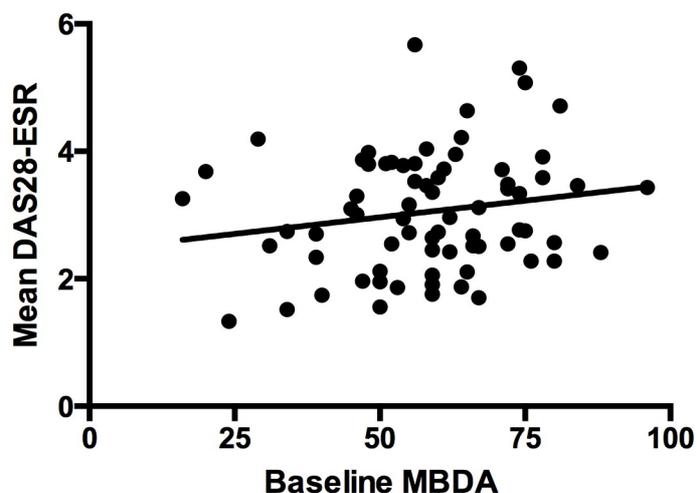
The degree of agreement between all of the disease activity measures discussed thus far was tested in the subset of 35 instances when there was concurrent DAS28-ESR, MBDA and MSUS disease activity assessment data available. The assessment methods were tested to determine how well they agreed on the need to escalate DMARD therapy based on existing escalation thresholds that differentiate between low and moderate disease activity (i.e. DAS28 >3.2, MBDA > 29, MSUS PD joint count  $\geq$  2). Separately, there were 4 instances (11%) when DAS28-ESR supported DMARD escalation, 29 instances (83%) when MBDA supported DMARD escalation and 17 instances (49%) when MSUS assessment supported DMARD escalation. There were 2 occasions (6%) when none of the assessment methods supported DMARD escalation, 17 occasions (49%) when only 1 assessment supported DMARD escalation, 15 occasions (43%) when 2 assessment methods concurrently supported DMARD escalation and 1 occasion (3%) when all 3 assessment methods agreed. On the 17 occasions that only 1 assessment method support DMARD escalation, 13 (76%) were because the MBDA test identified moderate disease activity (or higher) and 4 occasions (24%) were because MSUS assessment identified active disease. There were no occasions when DAS28-ESR alone supported DMARD escalation. On the 15 occasions that 2 assessment measures both supported DMARD escalation, 3 (20%) occurred because of agreement between DAS28-ESR and MBDA and 12 (80%) occurred because of agreement between MBDA and MSUS assessment. There were no instances when DAS28-ESR and MSUS assessments agreed

### **8.3.5 Relationship between MBDA score and subsequent disease activity**

To determine whether MBDA test values possessed any ability to predict subsequent disease course, the relationship between MBDA score and mean DAS28 was investigated. As before, mean DAS28-ESR between baseline and 12 months was used to represent disease course. This analysis relates to a sub-group of 72 participants, since participants were excluded if there were less than 10 DAS28-ESR data-points available between baseline and follow-up month 12. The mean DAS28-ESR 0-12 months was 3.05 (SD 0.93).

Overall, only weak relationships were identified between MBDA and subsequent disease course. Baseline MBDA score correlated very weakly with mean DAS28-ESR 0-12 months ( $r_s = 0.14$ ,  $p = 0.24$ ) and the scattergraph depicting the relationship revealed significant dispersal of data-points

(Graph 61). Similarly, there was no relationship identified between change in MBDA between baseline and 3 months and mean DAS28-ESR 0-12 months ( $r_s = 0.07$ ,  $p = 0.58$ ).



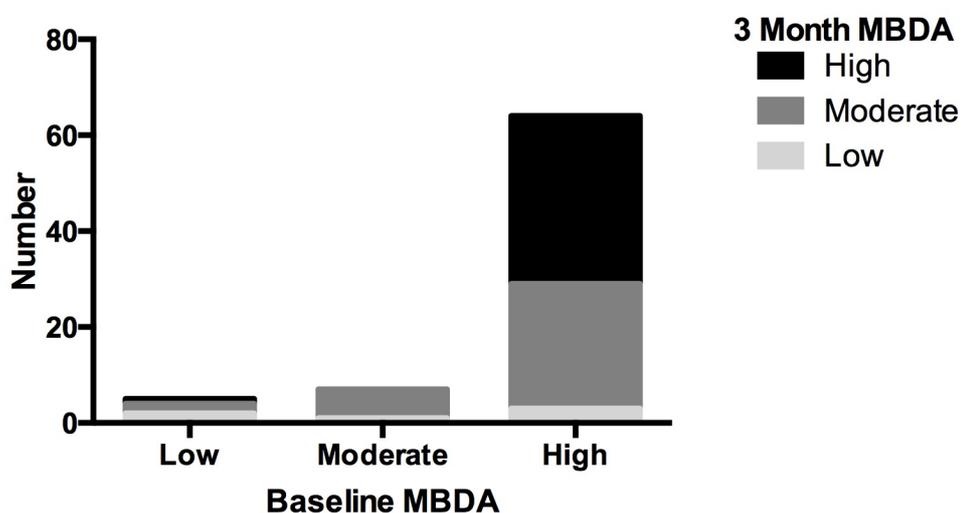
**Graph 61** – Scattergraph depicting relationship between baseline MBDA and mean DAS28-ESR between baseline and follow-up month 12

Subdividing participants into quartile groups of the mean DAS28-ESR 0-12 months did not identify significant differences in baseline MBDA between those with the highest and lowest disease burden between baseline and 12 months. The medians of the (mean) DAS28-ESR 0-12 months for each quartile group were very similar: 1st quartile 56 (IQ 45 – 66), 2<sup>nd</sup> quartile 61 (IQ 54 - 68), 3<sup>rd</sup> quartile 60 (IQ 46 – 73) and 4th quartile 57 (IQ 50 – 67). Furthermore, stepwise comparisons between each quartile group demonstrated there was no statistical difference in the baseline MBDA between any of the groups ( $p = 0.35 - 0.88$ ). Similarly, there were no statistically significant ( $p = 0.13 - 0.83$ ) differences in median DAS28 0-12 months when participants were subdivided into groups based upon quartiles of the baseline MBDA score; for the lowest quartile, the median DAS28-ESR 0-12 was 2.88 (IQ 1.91 – 3.71), for the 2<sup>nd</sup> quartile - 3.05 (IQ 2.03 – 3.81), for the 3<sup>rd</sup> quartile - 2.70 (IQ 2.44 – 3.62) and for the 4th quartile - 3.42 (IQ 2.57 – 3.76).

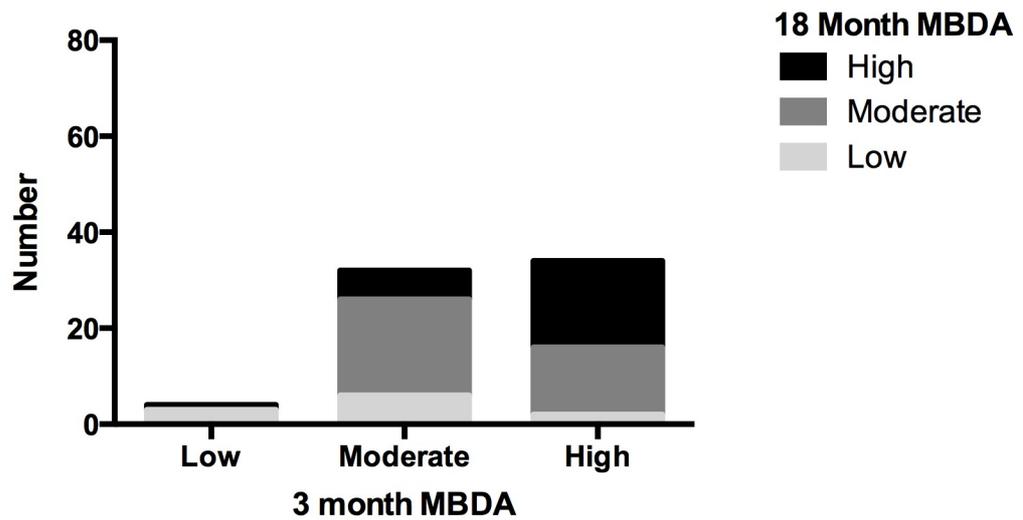
Stratifying participants by MBDA disease activity group at baseline and 3 months did not reveal any new associations between MBDA score and subsequent disease course. The analysis was limited because using the MBDA disease activity definition as a grouping variable created significance imbalance in the size of each group. At baseline, 4 participants were classified as low MBDA disease activity, 6 were classified as moderate MBDA disease activity and 62 were classified as high MBDA disease activity. Each participant's mean DAS28 between baseline and 12 months was used as an estimate of overall inflammatory disease activity. A specific pattern in the overall inflammatory exposure could not be identified. For the low baseline MBDA group, the median of the (mean) DAS28-ESR 0-12 months was 3.47 (IQ 1.8 - 4.1), for the moderate baseline MBDA group - 2.43 (IQ 1.69 – 2.71) and for the high MBDA group - 3.11 (IQ 2.44 – 3.79). Tentative between group comparisons revealed a statistically significant difference between the moderate and high baseline MBDA groups ( $p = 0.02$ ) only. Given the imbalance in group sizes, further statistical comparisons were not conducted. Three month MBDA exhibited a similar relationship to subsequent disease course (data not shown).

Participants categorised as high and moderate MBDA disease activity at baseline did experience a significantly greater change in DAS28 between baseline and 12 months than participants classified as low MBDA disease activity. The median change in DAS28-ESR between baseline and follow-up month 12 was -0.88 (IQ -1.67, -0.32) in the low MBDA group, -2.45 (IQ -2.73, -2.08) in the moderate MBDA group and -2.57 (IQ -3.67, -1.52) in the high MBDA group. Allowing for notable differences in group sizes, there was a statistically significant difference demonstrated in the change in DAS28-ESR between the low and moderate baseline MBDA groups ( $p = 0.016$ ) and low and high baseline MBDA groups ( $p = 0.014$ ) but not between the moderate and high MBDA groups ( $p = 0.81$ ).

Stepwise comparisons on how MBDA disease activity states at 3 months and 18 months compared to the preceding MBDA disease activity state suggested that a proportion of patients classified as high MBDA disease activity at either baseline or 3 months continued to be classified as moderate or high MBDA disease activity during follow-up assessments (Graphs 62 and 63). At baseline, 64 participants were classified as MBDA high disease activity. After 3 months of DMARD monotherapy, 3 participants (5%) were classified as low MBDA disease activity, 26 (41%) were now classified as moderate MBDA disease activity and 35 (55%) were still classified as high disease activity (Graph 62). Similarly, of the participants classified as MBDA high disease activity at 3 months, 2 (6%) were classified as low MBDA disease activity at 18 months, 14 (41%) were classified as moderate MBDA disease activity and 18 (53%) were still classified as high disease activity (Graph 63)



**Graph 62** – Outcome of 3 month MBDA disease activity categorisation based upon baseline MBDA disease activity state



**Graph 63** – Outcome of 18 month MBDA disease activity categorisation based upon 3 month MBDA disease activity state

### Multiple regression analyses incorporating MBDA

To determine whether baseline MBDA was an independent predictor of disease course, the previously described multiple regression analyses (Section 6.8) were rerun using MBDA as an independent (predictor) variable instead of baseline CRP for the subset of participants who contributed to the MBDA analysis. All other independent variables were unchanged

*Change DAS44 0 – 18 months*

$R^2 = 0.285$ ,  $F = 15.2$ ,  $p < 0.001$

Predictor Variable	Beta Co-efficient	p
Baseline DAS44	-0.39	<0.001
Baseline MBDA	-0.25	0.017

(Randomisation group, gender, anti-CCP antibody status, smoking status, age and symptom duration were not significant predictors in this model)

*Mean DAS44 0-18 months*

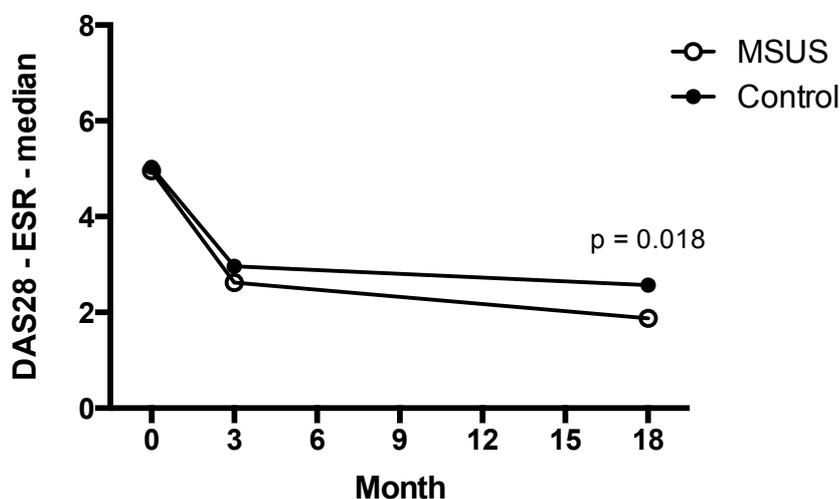
$R^2 = 0.31$ ,  $F = 16.7$ ,  $p < 0.001$

Predictor Variable	Beta Co-efficient	p
Baseline DAS44	0.59	<0.001
Baseline MBDA	-0.21	0.049

(Randomisation group, gender, anti-CCP antibody status, smoking status, age and symptom duration were not significant predictors in this model)

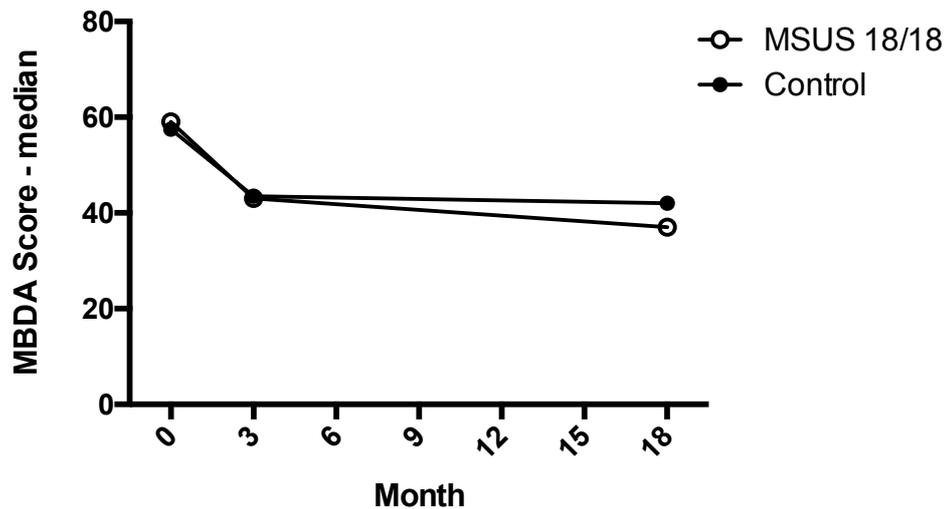
### 8.3.6 Change in MBDA score by randomisation group

For the MBDA analysis results described so far, all participants have been pooled together, regardless of their randomisation group allocation. Since the assessment protocols will have led to each group receiving differing intensities of DMARD therapy it is possible that this may have been reflected in differences in the change on MBDA result over the follow-up period. Therefore, participants were reformed into comparator groups based upon their original randomisation allocation; 38 participants were included within the DAS28 assessment group and 38 were included within the MSUS assessment group. For each group, the median DAS28-ESR was similar at baseline (DAS28 group 5.03 vs MSUS group 5.00), follow-up month 3 (DAS28 group 2.96 vs 2.63) and follow-up month 18 (DAS28 group 2.57 vs MSUS group 1.88) (Graph 64). There were no significant between group differences evident in the DAS28-ESR at either baseline and month 3 ( $p = 0.98$  and  $0.45$  respectively) though the difference was statistically significant at month 18 ( $p = 0.028$ ).



**Graph 64** – Median DAS28-ESR by randomisation group for participants included within the MBDA analysis  
(\*  $p = 0.035$  – no other comparisons were significant)

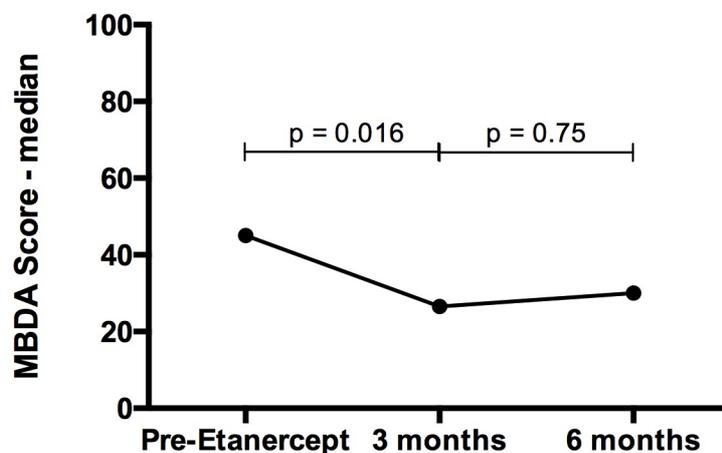
Overall there were no statistically significant differences evident in the MBDA scores between either randomisation group at any of the time points and the MBDA scores appeared to follow a very similar pattern (Graph 65). Indeed, median MBDA scores were very similar for both groups at baseline and follow-up month 3 (57.5 vs 59 and 43.5 vs 43 respectively). At follow-up month 18, the median MBDA score was slightly lower in the MSUS assessment group (DAS28 group 42 vs MSUS group 37). There were no significant between group differences identified in the mean change in MBDA between either baseline and 18 months (mean change MBDA: DAS28 group -15.7 vs MSUS group -20.9,  $p = 0.22$ ) or follow-up months 3 and 18 (mean change MBDA: -5.1 vs -7.4,  $p = 0.54$ ).



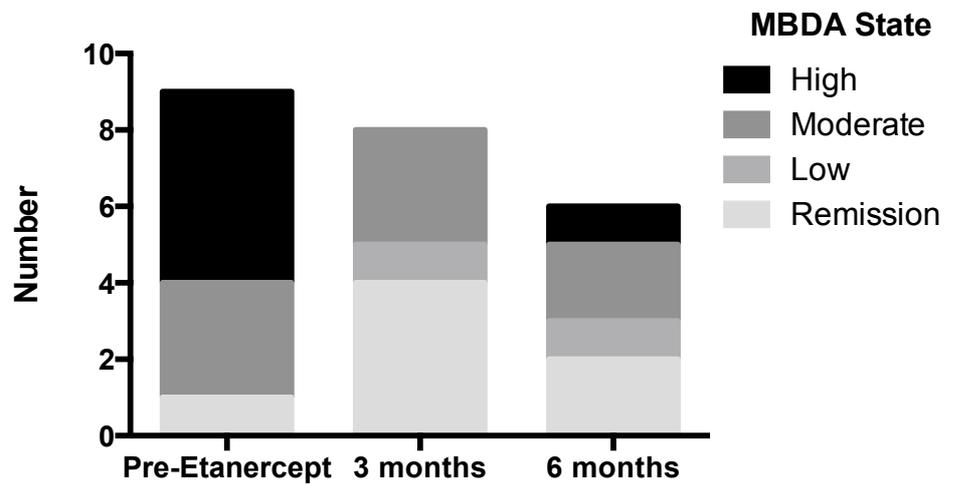
**Graph 65** – Median MBDA score by randomisation group at baseline, 3 months and 18 months. There were no statistically significant between group differences

### 8.3.7 Change in MBDA score in response to anti-TNF $\alpha$ blocking therapy

A small subset of participants commenced etanercept during the follow-up period and provided additional serum samples for analysis. Nine serum samples corresponded to participants commencing etanercept, 8 samples corresponded to 3 months of etanercept and 6 samples corresponded to 6 months (i.e. completion) of etanercept. Overall, commencement of etanercept coincided with a further decrease in MBDA (Graph 66). The median MBDA scores at each time point were: commencement of etanercept = 45 (IQ 34 – 61); 3 months of etanercept = 27 (IQ 16 – 36) and 6 months of etanercept = 30 (IQ 16 – 40). There was a statistically significant difference in median MBDA scores between commencement and month 3 of etanercept ( $p = 0.016$ ), but not between commencement and month 6 ( $p = 0.125$ ) nor month 3 and month 6 ( $p = 0.75$ ). The distribution of MBDA defined disease activity states in relation to etanercept is depicted in Graph 67. Overall, etanercept was associated with a reduction in the number of participants with high MBDA disease activity and increase in the number of participants with low MBDA disease activity or MBDA remission. However, given the relatively small group sizes, further statistical comparisons have not been performed.



**Graph 66** – Median MBDA score in relation to etanercept therapy



**Graph 67** – Distribution of MBDA-defined disease activity in relation to etanercept therapy

## 8.4 Discussion

The MBDA analysis described in this section continues the central theme of this research; namely, how do novel methods of global RA disease activity compare to established measures, such as the DAS28. Indeed, the MBDA test represents several steps further along in the development of a biomarker based objective measure of global disease activity than might be possible with the outputs of the gene expression analysis described in Chapter 7. Therefore, of the assessment methods tested in this study, the MBDA test is perhaps the one closest to being used outwith a research environment and in general clinical practice.

From the results of this analysis it is possible to make the following conclusions about the possible role of the MBDA test as a disease activity measure in RA:

1. The MBDA score falls following commencement of DMARD therapy and the steepest rate of change occurs in the months immediately following commencement of therapy.
2. MBDA at a single time point demonstrates a moderate positive correlation with DAS28-ESR at the same time point ( $r_s = 0.58$ ). The change in MBDA between two time points has a moderate positive correlation with the change in DAS28-ESR over the same time period ( $r_s = 0.56$ )
3. There is significant disagreement evident between classification of disease activity state by DAS28-ESR and MBDA. For moderate and low DAS28-ESR states, MBDA classifies the majority of patients as a higher disease activity state
4. The degree of agreement between DAS28-ESR and MBDA disease activity classification increases when the measures are used as dichotomous descriptors of disease activity (e.g. high/moderate disease activity vs low/remission)
5. A substantial number of instances that MBDA classifies as either high or moderate disease activity are not associated with MSUS evidence of active synovitis
6. Baseline MBDA score may be an independent predictor of change in DAS44 and mean DAS44 up until 18 months of follow-up

7. MBDA score and disease activity classification improves following commencement of etanercept therapy

*MBDA as a disease activity measure*

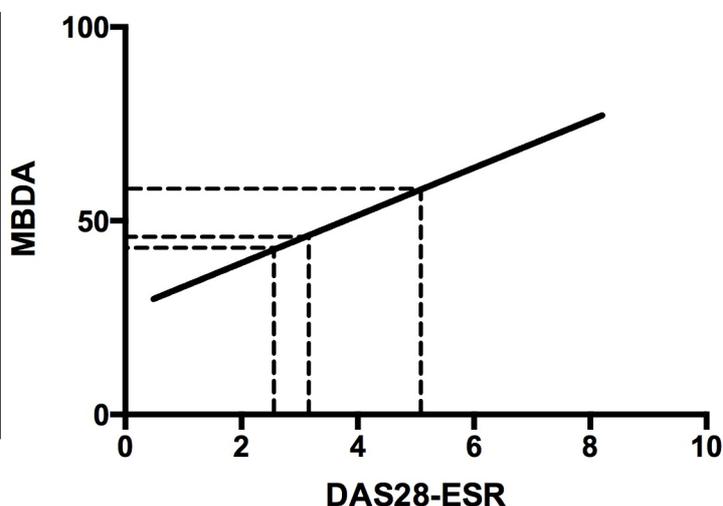
The correlation of the pooled results between DAS28-ESR and MBDA at a single time point ( $r_s = 0.58$ ), and the correlation between the changes in DAS28-ESR and MBDA over time ( $r_s = 0.56$ ), reported by this study are similar to those reported in previous studies. In the original validation paper, Curtis et al reported a moderate-strong positive correlation between the two measures at a single time point ( $r_p = 0.57$ ) and a moderate positive correlation between the change in DAS28-CRP and MBDA after 6-12 weeks ( $r_s = 0.51$ ,  $p < 0.001$ ) (390). In a subgroup analysis of the BeST study, Hirata et al reported a moderate-strong positive correlation between DAS28-ESR and MBDA at a single time point ( $r_s = 0.66$ ,  $p < 0.0001$ ) and a moderate positive correlation between the change in DAS28-ESR and MBDA after 12 months of treatment ( $r_p = 0.55$ ,  $p < 0.0001$ ) (393). In a subgroup analysis of the CAMERA study, Bakker et al reported a strong positive correlation between DAS28-CRP and MBDA at a single time point ( $r_p = 0.72$ ,  $p < 0.001$ ) (396). However, whilst Bakker et al have reported that patients treated using an intensive DMARD escalation regimen demonstrate a greater improvement in MBDA score, the relationship between the change in DAS28-CRP and MBDA scores is not specifically described.

Taken together, the consensus of the available results seems to suggest that the MBDA test shares a positive, linear relationship with DAS28-ESR and DAS28-CRP. That is, they represent similar static and dynamic changes in disease activity, even though their output values are not numerically comparable. Therefore, if the MBDA measures and reacts to disease activity in a similar manner to DAS28-ESR and DAS28-CRP it is feasible to presume that it might also be used as an alternative (though not interchangeable) disease activity measure. However, whilst the quoted correlation coefficients between static and dynamic DAS28-ESR, DAS28-CRP and MBDA values are consistently positive and statistically significant, it is clear that there is frequent disagreement between how each measure categorises disease activity in individual patients (Graph 57). Indeed, there were frequent occasions when the MBDA test had classified disease activity at a higher level than suggested by the DAS28-ESR (Graph 58) and there were a number of occasions when changes in the MBDA failed to reflect the same degree of improvement in disease activity as suggested by the DAS28-ESR. Interestingly, the highest degree of agreement between DAS28-ESR and MBDA in disease activity classification was recorded during the baseline assessments (Graph 57) but declined significantly during subsequent assessments. It is unlikely that this can be wholly explained by the influence of immunomodulatory therapy on the expression levels of the independent biomarkers that contribute to the MBDA calculations since immunomodulatory agents should cause a decline in the expression levels of pro-inflammatory biomarkers, resulting in lower final MBDA scores and, if anything, categorisation of disease activity at a lower level.

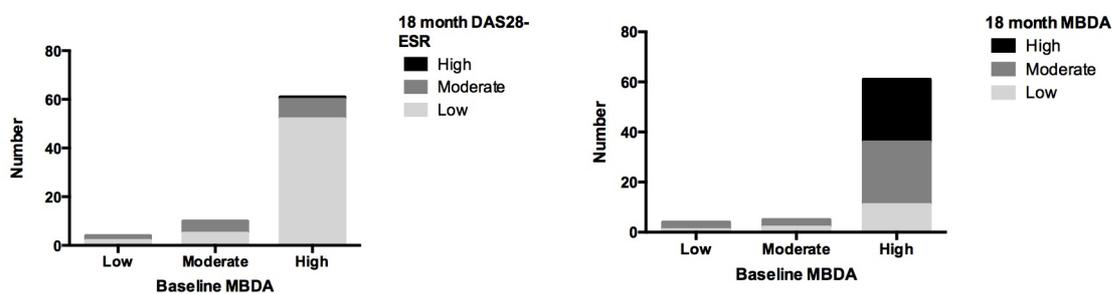
The MBDA is calculated using formulae that are derived from the DAS28-CRP. Whilst this leads to the DAS28-CRP and MBDA scores having a consistently positive relationship, it is clear from the correlation analyses that the relationship is not perfectly proportional. Hence, each measure

may not change by the same proportion in response to changes in overall disease activity. Indeed, interpolations made using the linear regression of the correlation between DAS28-ESR and MBDA (Graph 68 and Table 54) estimate that the corresponding MBDA thresholds for each disease activity state for this cohort are somewhat higher than in the previously published studies. Similarly, the MBDA score remained elevated in the majority of participants after 18 months of treatment despite them having attained DAS28-ESR LDAS. At baseline, 61 participants were categorised as high disease activity by the MBDA test; after 18 months of treatment, 52 of these participants had attained DAS28-ESR LDAS. Conversely, after 18 months, 25 participants each continued to be categorised as moderate and high MBDA disease activity (Graphs 69A+B). Thus, either the MBDA continues to detect on-going subclinical synovitis in patients with low disease activity, its rate of change over 18 months is not proportionate to the corresponding change in DAS28-ESR and/or there is a subset of RA patients whose MBDA score remains elevated despite an adequate clinical response to DMARD therapy.

DAS28-ESR	Corresponding MBDA
Remission DAS28 <2.6	42
Low DAS28 <3.2	45
High DAS28 >5.1	58



**Table 56 and Graph 68** – MBDA disease activity definition thresholds for this cohort – estimated by interpolating linear regression of correlation between all DAS28-ESR and MBDA (x axis intercepts correspond to DAS28 = 2.6, DAS28 = 3.2, DAS28 = 5.1)



**Graphs 69A + 69B** – Relationship of baseline MBDA disease activity to 18 month DAS28-ESR and MBDA disease activity grouping

Comparisons between DAS28-ESR and MBDA disease activity categorisation for this study have consistently shown that MBDA frequently categorises disease activity at a higher level than the corresponding DAS28-ESR (Graph 58). Most notably, 82% of the instances of DAS28-ESR LDAS were classified by the MBDA as either moderate or high disease activity. This may have

treatment implications, since identifying high or moderate disease activity encourages DMARD escalation, whereas identifying low disease activity does not. A similar disagreement was not identified in relation to high DAS28-ESR disease activity since virtually all instances (95%) of DAS28-ESR high disease activity were also categorised as high MBDA disease activity. This raises several interesting points. Either: 1. MBDA scores are related to RA pathogenesis and are independent of disease activity, 2. MBDA is more sensitive than DAS28-ESR at low levels of disease activity and/or 3. MBDA is less specific than DAS28-ESR. It seems likely that the MBDA score is responsive to changes in disease activity since this study, and several of the previous validation studies, have demonstrated that the MBDA score correlates positively with existing measures of global disease activity (such as DAS28-CRP, DAS28-ESR, SDAI) across their whole range of values and that changes in global disease activity measures are matched by corresponding changes in MBDA (390,393,396). So far, the MBDA has been validated against existing global disease activity measures that rely upon clinical detection of painful and swollen joints. However, as previously discussed, there remains concern about the sensitivity of these composite measures to detect subclinical active synovitis and their specificity in patients with multiple causes of joint pain. Therefore, it may be difficult to estimate the specificity of a new measure (MBDA) if the existing validation measure (DAS28) against which it is being compared is less sensitive at low disease activity levels (i.e. the MBDA will appear to generate a high proportion of false positive results)

The insensitivity of composite measures to low level active synovitis may partly explain why the MBDA consistently reported a higher level of disease activity than the corresponding DAS28-ESR measure; that is, MBDA may 'reveal' subclinical disease when clinical examination is otherwise unremarkable. By extension, if composite disease activity measures are insensitive to low levels of active synovitis, it may be more appropriate to test the performance of the MBDA against modern imaging techniques (such as MRI or MSUS). The availability of corresponding MSUS disease activity findings for a subset of occasions has allowed the relationship between MBDA and subclinical synovitis to be partially investigated. Overall, the available results suggest that, within the limitations of the MSUS joint set and definitions of active disease, high and moderate MBDA scores are not always associated with ultrasonographic synovitis (Graph 60 and Table 55). Indeed, only 42% of the occasions classified as high disease activity by MBDA demonstrated active synovitis during MSUS assessment. Similarly, 41% of the occasions of moderate MBDA disease activity and 67% of the occasions of low MBDA disease activity also had evidence of active synovitis during MSUS assessment. The number of instances when any joint exhibited PD signal (i.e. PD joint count  $\geq 1$ ) were notably higher for all MBDA disease activity groups (high 75%, moderate 82%, low 67%); hence it may be that the elevation of MBDA is related to the presence of any active synovitis rather than the total number of joints that are affected. Nevertheless, the frequent absence of ultrasonographic evidence of synovitis in participants with moderate and high MBDA scores also raises additional points to consider. Either: 1. MBDA is even more sensitive than MSUS at detecting synovitis, 2. MBDA measures aspects of systemic inflammation that are not visible at the joint and/or 3. elevation of MBDA is related to an additional external factor and therefore may not be specific to RA disease activity. The present results represent a relatively small cohort of participants so should be considered preliminary at best;

however, the absence of ultrasonographic synovitis in participants with high and moderate MBDA scores inevitably does raise questions about the specificity of the MBDA. Additional analyses will be possible once the results of the 18 month MRI RAMRIS assessments are available. Indeed, a further analysis will be performed that correlates the 18 month MBDA score to the presence (or absence) of MRI synovitis and bone marrow oedema. However, once again, it is unlikely that this additional analysis will be large enough to fully define the specificity of the MBDA test. During the original development studies, Centola et al did attempt to determine the influence of potential confounders by comparing the MBDA scores between RA patients with and without each of several common comorbidities (hypertension, osteoarthritis, osteoporosis, diabetes mellitus, asthma) (383). Overall, the presence of an additional comorbidity was not associated with a significant difference in the measured MBDA scores. However, the list of confounders is relatively limited and, other than asthma, does not contain any other inflammatory conditions (e.g. intercurrent infection, psoriasis, inflammatory bowel disease) that could also feasibly influence the expression of pro-inflammatory markers and the final MBDA score.

At present, it may be difficult to incorporate the MBDA test into routine clinical practice, until its true sensitivity and specificity in relation to RA activity have been clarified. The balance between sensitivity and specificity will be crucial to the long term clinical usefulness of the MBDA test. An objective and sensitive additional blood marker of joint inflammation could provide extremely useful additional disease activity information that helps target DMARD treatment against persistent active synovitis in patients who no longer demonstrate clear clinical synovitis. However, for the test to be clinically useful its specificity for active synovitis will need to be high enough to prevent unnecessary, and potentially toxic, DMARD escalation in patients who have no on going inflammatory joint disease. Eventually a prospective study will be required to demonstrate that DMARD escalation regimens targeting MBDA low disease activity, and/or remission, produce significantly better outcomes than current, DAS28 and SDAI based regimens without also causing additional adverse effects.

#### *The predictive properties of the MBDA*

So far, several studies have addressed the potential predictive ability of MBDA in relation to radiographic progression, though their findings have not been consistent. In a subgroup analysis of the BeST study, the baseline MBDA score was not predictive of radiographic progression after 2 years (392). By contrast, in the SWEFOT study, there was significant positive correlation between baseline MBDA score and the change in radiographic damage score after 1 year (399). Similarly, Van-der-Helm-van-Mil et al have demonstrated that the persistent elevation of MBDA score in some patients may be pathogenetically relevant since it is associated with an increased risk of radiographic progression after 1 year (394). Fewer patients who attained MBDA remission experienced radiographic progression compared to patients who'd attained either DAS28-CRP and/or ACR-EULAR Boolean remission (proportion of radiographic progression after 12 months = 7% vs 20% vs 17% respectively). Furthermore, the risk of radiographic progression was proportional to the MBDA disease activity grouping and, importantly, the subset of patients with DAS28-CRP remission but high MBDA disease activity were at a significantly increased risk of radiographic progression (relative risk 2.28; 95CI 1.13, 3.68) compared to all patients in DAS28-

CRP remission. The implication being that persistent elevation of MBDA is associated with active disease because it is also associated with progressive joint damage.

The results of this study do suggest that MBDA may be an independent predictor of disease course. Linear correlations suggested that baseline and 3 month values of MBDA had a very weak positive correlation with subsequent disease course. However, multiple regression analysis did suggest that baseline MBDA was a significant, independent predictor of both change in DAS44 and mean DAS44 between baseline and 18 months. However, it appeared to have a similar influence on the statistical model as the cheaper (and widely available) baseline CRP value. The ability of baseline MBDA to predict future radiographic progression remains to be determined for this study

## **9. Final Discussion and Considerations**

The TaSER study was originally devised to investigate whether using a MSUS measure of global disease activity in early RA would improve patients short-medium term clinical, functional and radiological outcomes. Since blood samples were collected at set time points throughout each participant's follow-up period it has been possible to broaden the scope of the study to consider in general the potential value of several alternative methods of assessing global disease activity. Interestingly, the results have been somewhat mixed: musculoskeletal ultrasound provided additional disease activity findings (and often altered treatment decisions) during a notable proportion of assessments but was not associated with significantly better clinical outcomes; peripheral blood gene expression analysis did not have any relationship with either RA disease phenotype or disease activity state; however, an inflammatory protein based MBDA score was more sensitive than DAS28-ESR and MSUS assessments but may be non-specific and is yet to be tested in an interventional treatment strategy study. Taken together, these results start to suggest that novel, technology dependent, methods of assessing (and monitoring) global disease activity may provide minimal additional benefit to existing, largely clinical, assessment methods.

The wide spread use of quantitative, composite disease activity scores is now accepted as having improved the sensitivity and objectivity of clinical assessments of global disease activity and contributes significantly to the efficacy of current early RA treatment strategies (124,170,350). Indeed, several interventional treatment strategy studies have associated systematic assessment of global disease activity with improved clinical and radiological outcomes (118,119,128). Though whether the measured improvements in outcomes were related directly to the method of disease activity assessment, or simply reflected the impact of using more aggressive therapy, will be difficult to disentangle since they comprise one component of a complex intervention that also includes early DMARD initiation, aggressive intensification of treatment, early (often generous) use of corticosteroids and early multidisciplinary team input. Regardless, formalised assessment of global disease activity has clearly enabled clinicians to gain better short-medium outcomes and is now accepted as a routine care standard. However, evidence from recent imaging studies has started to suggest that current thresholds for LDAS and/or remission do not necessarily correspond to either inactive synovitis (121,150,400,401) nor an absent risk of disease progression (200,217). Therefore, there has been a mounting momentum of opinion that DMARD escalation strategies should target a stricter definition of remission rather than clinical remission, since existing composite disease activity assessment methods are relatively insensitive compared to modern imaging techniques (129). However, whether even tighter control of inflammatory joint disease is associated with significantly better outcomes, and an acceptable risk:benefit ratio, has not yet been demonstrated in a randomised clinical trial. Equally, the additional benefits associated with novel methods of disease activity assessment will need to be striking in order to be distinguishable at a group level from the benefits that can be achieved using existing treatment strategies and disease assessment methods.

The TaSER study is the first clinical trial to report the influence of MSUS disease activity assessment on any outcome measure in early RA. Whilst the OMERACT Imaging Group continues to debate which joints will comprise the limited joint set of the Global Ultrasound

Synovitis Score (294), the results reported herein provide the first indication that regularly assessing global disease activity using MSUS may not significantly improve clinician's ability to monitor and treat early RA, at least in the short-medium term. Therefore, the wide spread up take and use of MSUS by rheumatologists, at least for disease activity monitoring purposes, is not supported. Indeed, DAS28-driven DMARD escalation was still associated with very good, and similar, clinical and laboratory outcomes to MSUS-driven DMARD escalation without having to use quite such aggressive combinations of DMARDs and biologics, nor spend additional time during busy clinics conducting MSUS assessments.

During DAS28 LDAS and remission states, MSUS did frequently identify evidence of on going disease activity, but aiming to achieve imaging remission by suppressing this synovitis does not appear to have produced significantly better short-medium term clinical outcomes. However, whether this approach is associated with lesser rates of damage progression on plain x-rays and MRI images and/or better medium-long term clinical, functional and radiographic outcomes remains unanswered. It is important to consider that the current conclusions are based upon an incomplete dataset. The potential impact of MSUS-driven DMARD escalation on adverse event rates and DMARD intensity has not yet been specifically reported. It may become even harder to justify the additional costs (equipment, treatment burden and financial cost) of regular MSUS disease activity assessments if MSUS group participants also appear to experience a higher rate of treatment associated adverse events for no additional benefit. Similarly, a number of outstanding factors need to be clarified before concluding that MSUS assessment conferred no additional benefits on the MSUS group. Firstly, the MSUS assessment group did experience incremental increases in DAS44 remission and EULAR good response rates over the follow-up period and did also demonstrate a higher DAS44 remission rate at the 18 month follow-up point. It is possible that these findings reflect emerging divergence of outcomes between the assessment groups that would have become more apparent over a longer follow-up period. Secondly, the MSUS assessment group may yet demonstrate lesser rates of progression in MRI and plain x-ray measures of joint damage. If this is the case, the lack of clinical benefit may be counter-balanced by evidence that aggressive early suppression of active synovitis does attenuate (if not halt) progression of joint damage that should be associated with improved medium-long term clinical and functional outcomes. These additional benefits may not have been immediately apparent during the relatively restricted initial follow-up period.

Interestingly, during instances of DAS28 moderate disease activity, MSUS disease activity assessment did still contribute useful, additional information regarding disease activity. In fact, MSUS assessment may yet prove most useful in the assessment of symptomatic patients with equivocal, or absent, examination findings. It may not be possible to use MSUS disease activity assessment to improve the treatment outcomes of RA patients with DAS28 LDAS and/or remission but it may be possible to use MSUS to ensure that patients with DAS28 moderate disease activity but minimal clinical synovitis are considered for the most appropriate therapeutic intervention, especially if there is minimal clinical evidence of synovitis. In patients with evident clinical synovitis MSUS is unlikely to add any additional information to the assessment of disease activity unless there is concern that joint swelling represents chronic synovial hypertrophy rather

than active synovitis. Similarly, patients with DAS28 moderate disease activity but minimal clinical synovitis will benefit from MSUS assessment if it differentiates between patients with active subclinical synovitis, who may benefit from DMARD escalation, and patients with no active synovitis, who may not benefit from DMARD escalation, but may benefit from an alternative treatment approach. In all these instances, MSUS findings will have contributed to individual patients being offered the most appropriate treatment for their specific needs. Those with active synovitis can still be considered for DMARD escalation; whereas those without active synovitis can be spared the risks of unnecessary DMARD escalation but can be considered for alternative treatment approaches instead. Admittedly, it appears that MSUS assessment will influence treatment decisions for instances of DAS28 moderate disease activity much less often than for instances of DAS28 LDAS. Therefore, whilst it may remain a useful additional assessment method for symptomatic patients with equivocal clinical examination findings it may prove difficult to prove the efficacy in an adequately powered clinical trial.

This study has found conflicting evidence about the potential value of using biomarker panels to assess RA phenotype and disease activity. It does not appear that specific patterns within peripheral blood gene expression profiles associate with particular RA disease characteristics nor extremes of RA disease activity. Therefore, it is unlikely that, in its present guise, peripheral blood gene expression profile assessment will be able to add useful additional information to current methods of describing either phenotype and/or disease activity. This conclusion may appear excessively pessimistic and in due course it may well be shown that particular gene expression patterns do consistently associate with particular phenotypic features. The comparator groups used during the gene expression analysis were deliberately devised to reflect common clinical scenarios so that the clinical relevance of any positive findings would be immediately evident. However, these common scenarios were not associated with clear evidence of differential gene expression. Further analysis is required to determine whether specific phenotypic sub groups do exhibit evidence of differential gene expression (e.g. restricted to autoantibody positive patients, limited to a single drug therapy); however, as the cohort is split into smaller and smaller subgroups the potential for skew within the statistical analysis processes increases and the clinical relevance of any positive findings diminishes. Ultimately, specific gene expression profiles that relate to very strictly defined clinical scenarios may prove too costly to develop and commercially unviable. It is possible that gene expression profiles may provide important pharmacogenomics information relating to an individual patient's risk of developing a drug related adverse effect (such as thiopurine methyltransferase genotype in relation to azathioprine (402)). However, it is unlikely that the TaSER cohort was large enough, nor had a high enough event rate, to have sufficient power for this to be analysed adequately. Several studies have reported associations between tissue (usually peripheral blood or synovial tissue) gene expression profile and RA phenotype; however, in most cases these studies have been either relatively small and/or have not yet validated the findings in independent cohorts. There is perhaps a little more evidence to support an association between synovial tissue gene expression profile, disease activity and treatment response (236,241). However, since patients are required to undergo a synovial biopsy the clinical usefulness of synovial tissue analysis remains unclear. Altogether, the findings of the TaSER gene expression analysis are largely in step with the findings of other

published studies in RA (403,404). That is, almost all recent studies have described the association between gene expression profile and phenotype; however, until robust relationships between specific gene expression patterns and clinical phenotype are identified (and validated) it won't be possible to 'test' the clinical value of the profile in a clinical trial setting

The available results suggest that the inflammatory protein based MBDA test is perhaps closer to being adopted as an additional measure of disease activity than gene expression based assessments. Indeed, the MBDA test remains a very attractive alternative disease activity measure since it is relatively easy to obtain, fits with the understanding of RA as a systemic disease and removes any subjectivity related to patient reported outcomes and clinician interpretation of clinical findings. The findings of the TaSER study are very similar to previous comparisons between MBDA and DAS28 measured disease activity (383,390,392-394) and confirm that there is a significant positive correlation between DAS28ESR and MBDA at a single time point and a significant positive correlation in the degree of change of both measures over time. Therefore, it seems reasonable to conclude that the outputs of the DAS28 and MBDA are measuring similar disease states, even if they are not directly interchangeable. Even though the correlation analyses suggest that changes in DAS28 and MBDA are 'in step', comparison of how disease activity states are categorised has demonstrated disagreement between the different measures. The MBDA test tends to categorise disease activity at a higher level than the DAS28 and, in part, this suggests that the MBDA may be a more sensitive measure of very low levels of inflammatory joint disease than DAS28. However, over half (approximately 58%) of high and moderate MBDA assessments were not associated with MSUS evidence of active synovitis and there were a subset of participants who continued to exhibit high MBDA disease activity even though they had attained a good clinical response. Both these latter findings raise questions about the MBDA test's specificity in relation to the presence of active synovitis. In due course, being able to correlate 18 month MRI RAMRIS synovitis score to corresponding MBDA score will help clarify some of the issues relating to both sensitivity and specificity, though it must be accepted that the available MRI images will correspond to an even more limited joint set than the MSUS assessment. Once again, all of the published results merely describe the numerical relationship between DAS28 disease activity and MBDA; the potential value (and safety) of the MBDA as an alternative disease activity measure will not be fully apparent until it has been used to steer DMARD therapy in a controlled clinical trial setting. It is worth considering that persisting elevation of the MBDA score during DAS28 LDAS may be akin to the persistence of PD signal during MSUS assessment. This study has shown that both MBDA score and PD signal remain elevated in a significant subset of RA patients with DAS28 LDAS and previous studies have also shown that this is predictive of future radiographic progression (200,394). However, this study has also shown that attempting to fully suppress ultrasonographic evidence of active synovitis is not necessarily associated with improved clinical outcomes and it may be that there are similar findings from attempts to normalise MBDA using aggressive DMARD therapy.

Taken altogether the results of this study suggest that MSUS assessment and the MBDA test do provide additional disease activity information to the DAS28 in a subset of RA patients that could be used to modulate DMARD treatment decisions. However, so far, the available results do not

suggest that these novel methods necessarily lead to improved clinical outcomes. It may be that novel assessment methods are unable to improve upon the already clear benefits associated with DAS28 driven DMARD therapy and there will be little value in attempting to alter current practice if there are no additional benefits for patients. If MSUS-driven DMARD therapy is associated with improved radiological outcome, and/or better medium-long term outcomes, it may yet be possible to create an argument for incorporating some form of MSUS assessment into current practice. However, it is possible that many patients may find the associated increased treatment burden unacceptable if, symptomatically, they felt no different. Until the safety and efficacy of MSUS driven DMARD therapy has been fully described it is likely that DAS28 will remain the most commonly used method of assessing global disease activity in RA.

## Considerations

The research described herein has tried to address a number of related issues relating to assessment of RA disease activity. The structure of the TaSER study aimed to replicate routine clinical practice, deliver high quality care for participants as well as providing a structure for collection of clinical outcome data and tissue samples. Since this was a 'real world' setting and the assessment of disease activity and DMARD treatment protocol constituted a complex intervention there were a number of important factors that should be considered since they may have impacted upon the observed results of the clinical study and the findings of both the gene expression and MBDA test analyses.

### 9.1 The TaSER study

#### Length of follow-up period

Previous trials of different DMARD regimens in early RA have been able to demonstrate significant differences in outcome after 12 (120,123,143,147), 18 (118) and 24 (149) months. A follow-up period of 18 months was chosen since this fitted with the time available to conduct the study and previous, similarly designed studies (the TICORA study) had demonstrated very significant differences over a similar time period. Both groups received the same DMARD monotherapy for the first 3 months; therefore, the first point that randomisation group could have influenced DMARD treatment and future outcomes was actually follow-up month 3 and, in effect, the follow-up period is actually foreshortened to 15 months. Whilst there was no significant difference overall in ACR core set variables at any of the assessment time points the MSUS group did demonstrate an incremental increase in the proportion of participants achieving EULAR good responses, DAS44 and ACR/EULAR remission. Optimistically, all other factors appearing equal between the groups, this does suggest that a divergence of the disease activity measures may have become apparent had the follow-up period been longer.

MSUS disease activity assessments continued to identify active synovitis over the whole of the follow-up period. However, the majority of MSUS assessments did not identify active synovitis and did not lead to additional DMARD escalation. In fact, out of 414 MSUS assessments, only 107 (26%) fulfilled the proposed definition of active synovitis. Further, these 107 positive MSUS assessments comprise a small minority (14%) of the total number of assessments that the MSUS group underwent between follow-up months 3 and 18. Hence, the positive influence of MSUS disease activity assessment on DMARD escalation may have been too infrequent to produce a significant improvement in the measured outcomes over the duration of the follow-up period.

#### Group Size

The previously described sample size calculation (section 2.2.4) aimed to provide sufficient statistical power to detect a clinically significant difference in the mean change of DAS44 ( $\geq 1.1$ ). This was based on previously published studies that had demonstrated that the standard deviation of the fall in DAS44 was 0.7 (405). In fact, the standard deviation of the mean change in

DAS44 overall and for both groups was higher than 0.7 (Overall 1.49, DAS28 1.47, MSUS 1.51), suggesting that the research's sample size may not have been large enough to identify a clear between group difference. However, the 95% confidence intervals of the mean change in DAS44 for both groups overlap significantly (DAS28 -2.92, -2.10; MSUS -3.20, -2.33) and the boundaries of the 95% confidence interval (-0.84, 0.33) provided by the t test comparing the mean change in DAS44 between the groups are less than the minimally clinically significant difference. Considering both these facts, it is unlikely that increasing the group sizes would have substantially increased the chances of identifying a significant between group difference

The measured standard deviation of the mean change in DAS44 (1.49) is similar to that reported in the TICORA study (1.1-1.4). The TICORA study also based its sample size estimation of the previously reported standard deviation of DAS44 being 0.7. However, in both the TICORA and TaSER studies, the same research nurse (Sr Anne Stirling) conducted the DAS44 assessments and therefore it may have been more appropriate to use a standard deviation of the change in DAS44 which more closely matched her previously measured variability.

### **Baseline characteristics of assessment groups**

Overall, the randomisation process produced groups that were well matched for most baseline disease activity, inflammatory marker and functional ability variables. The minimisation process did lead to a numerical imbalance in the proportion of females in each group (DAS28 75% vs MSUS 61%,  $p=0.10$ ) which may have predisposed the DAS28 group to exhibiting a lesser initial treatment response, and a lesser rate of remission. The DAS28 group functioned as a control group, so a higher proportion of female participants should have made it easier to detect a between group difference since it may have predisposed the DAS28 group to exhibit a lesser treatment response and therefore higher disease activity measures. In fact, the available results suggest that the sex distribution is unlikely to have adversely influenced the observed results. At a whole cohort level and a group level, there was no between sex difference in DAS44 values at any time point, nor the mean change in DAS44 from baseline and multiple regression analyses did not suggest that gender was an independent predictor of either DAS44 or HAQ at 18 months. Therefore, it does not appear that an imbalance in gender distribution between the groups adversely affected the primary outcome. In line with previously reported results, male participants overall, and males within the DAS28 group, did return lower HAQ scores than females. However, comparisons between gender subgroups of the randomisation groups did not suggest that males or females in either randomisation group experienced significantly worse functional outcomes than same-sex members of the other randomisation group. In summary, in uncorrected analyses participant's gender did not appear to be associated with a significant difference in DAS44 levels at either a cohort or group level. Therefore it seems unlikely that the unequal allocation of female participants between the groups has caused a significant bias in the whole group outcomes.

### **Proposed MSUS definition of active disease**

As discussed in the preceding section, there remains no universally agreed definition of active synovitis which incorporates MSUS PD findings. Nor is there a universally agreed limited MSUS

joint set that is thought to accurately represent global disease activity. The proposed definition of active disease was pragmatically chosen to represent a level of synovitis that was sufficiently suggestive of polyarticular joint inflammation to justify DMARD escalation, but also minimised (but did not totally remove) the potential influence of PD artefact on disease activity perceptions. A higher PD joint count threshold for DMARD escalation would have further reduced the risk of misclassifying disease activity but may have restricted the impact of MSUS assessment on DMARD escalation. A lower PD joint count threshold would have led to a much higher rate of DMARD escalation within the MSUS assessment group, which might have been reflected as much clearer benefits over the follow-up period. Whilst, in a limited joint set a PD threshold of 1 for DMARD escalation would be more sensitive to the presence of active disease it is also more likely to be biased by the presence of PD artefact; such that, failure to differentiate PD artefact from true PD signal could lead to inappropriate DMARD escalation in patients who do not have active synovitis. Coincidentally, the performance characteristics of the proposed MSUS set, and a number of similar variations, are currently under investigation by the OMERACT Musculoskeletal Ultrasound Working Group to reach a consensus agreement on the best performing Global Synovitis Score (GloSS) (294)

The proposed indications for MSUS assessment aim to identify 2 opposing clinical scenarios and could have had competing influences on overall DMARD exposure in the MSUS group. During instances of DAS28 LDAS, MSUS assessment aims to identify subclinical synovitis with the intention of supporting DMARD escalation if active disease is confirmed. During instances of DAS28 moderate disease activity and minimal synovitis, MSUS assessment will support DMARD escalation if active disease is identified, but is more likely to prevent escalation in patients who would otherwise have qualified based upon the DAS28 findings. There were 92 instances of DAS28 LDAS when additional DMARD escalation was supported and there were 30 instances of moderate disease activity when further DMARD escalation was prevented by MSUS assessment. So, the net overall impact of MSUS assessment was to support additional DMARD escalation in the MSUS group. Equally, the proposed MSUS examination is confined to a limited joint set and may, in some instances of DAS28 moderate disease activity, have missed ultrasonographic synovitis by not examining the relevant joint areas. Thus, the positive impact of DMARD escalation in DAS28 LDAS states may have been tempered somewhat by restriction of DMARD escalation during instances of DAS28 moderate disease activity but minimal clinical synovitis if the MSUS assessment did not include the correct joint(s). Either way, MSUS examination failed to identify PD signal in only 30 instances of DAS28 moderate disease activity but minimal clinical synovitis. This accounts for a small proportion (7%) of all the MSUS assessments and an even smaller proportion of the total number of consultations attended by the MSUS group (4%). So, whilst failure to identify ultrasonographic synovitis may have restricted the impact of DMARD therapy, it is unlikely to have occurred sufficiently often to have significantly altered the measured outcomes.

### **Choice and performance of clinical outcome measures**

Most of the ACR core set measures (excepting HAQ and laboratory measures of the acute phase response) rely on clinical assessments to identify the presence of swollen and/or tender joints.

However, this thesis has argued that clinical assessment alone may be insufficiently sensitive to accurately describe very low levels of disease activity, especially when sub clinical synovitis may only be detected by advanced imaging techniques. Since the MSUS intervention focussed on the detection of clinically undetectable disease, the chosen DAS44 outcome measure may have been insufficiently sensitive to differentiate between those participants with true remission and those with persisting subclinical synovitis. Thus, participants in both groups would be categorised by DAS44 as LDAS and/or remission when in fact some still had active, subclinical disease. Despite these concerns, the DAS44 has clearly shown that both assessment groups experienced significant improvements in clinical disease activity over the follow-up period. Whilst there was no consistent difference identified in the degree of clinical treatment response it is possible that the use of MSUS guided DMARD escalation, which aimed to abrogate all PD activity within the limited joint set, was associated with lesser rates of persisting PD signal, and therefore a lesser risk of subsequent flare and radiographic progression. The DAS44 will not represent this difference, nor is there MSUS data available for comparison with the DAS28 assessment group. However, the awaited analysis of the MRI RAMRIS outcomes may demonstrate that MSUS-guided DMARD therapy was associated with greater improvements in MRI measures of erosion, synovitis and bone marrow oedema (see below)

To maximise validity, all of the main clinical outcomes for this study were measured independently by an experienced metrologist who was blinded to each participant's randomisation group and DMARD regimen. Therefore, the clinical outcomes for this research are considered to be single blinded with the potential for inter-observer variability reduced through the use of a single metrologist. Feasibly, the ability of the study to identify a between group difference may have been restricted if either group had contained a higher proportion of participants with overlapping causes of joint pain that contributed to elevated DAS44 assessments. However, based on analyses of the individual DAS44 components this seems extremely unlikely. The summary statistics and distribution of all DAS44 components were virtually identical at each of the assessment time points and it is therefore unlikely that there was a significant sub-set in either assessment group that contributed to skewed DAS44 measurements.

## **Radiographic outcomes**

At the time of writing the extraction of the baseline and 18 month DICOM files for the plain x-ray and MRI images from NHSGGC PACS system had just been completed. Once all the files have been anonymised, relabelled and catalogued, they will be forwarded to the collaborating radiologists for formal grading. Each image will be graded independently by two radiologists in the manner described in Section 2.6.5 and it will thus become possible to report whether MSUS guided DMARD escalation is associated with slower rates of radiographic damage progression.

Even if plain x-ray assessment of joint damage does not show any significant differences in rates of progression between the assessment groups, it is possible that differences in MRI measures of disease activity and damage will be present. At a single time point the three components of the MRI RAMRIS scoring system provide separate markers of i. current disease activity (synovitis), ii. joint damage up until that point (erosion) and iii. risk of future erosive damage (bone marrow

oedema) (303,406). However, conjecturing, it is possible that the apparent increases in ACR/EULAR remission, DAS44 remission and EULAR good responses rates observed in the MSUS group will also be associated with more favourable underlying MRI findings. It will be particularly interesting to determine whether MSUS guided DMARD escalation is associated with lesser total RAMRIS synovitis (i.e. a better treatment response as measured by MRI) and/or bone marrow oedema scores (i.e. a lower risk of future erosive progression) than the DAS28 group.

### **Adverse events rate**

Since there appears to be only a modest additional clinical benefit from using MSUS disease activity assessment it is particularly important to ensure that the greater number of DMARD escalation steps is not also associated with an increased risk of DMARD related adverse effects. Adverse event data were collected prospectively and was categorised according to nature, severity and likely relationship to DMARD therapy. At present, the adverse event data is being validated by the biostatisticians at the Robertson Centre for Biostatistics. In due course, an adverse events analysis will be conducted. Individual adverse events will be classified according to severity (as described in Section 2.6.6) and category (e.g. gastrointestinal, haematological, respiratory etc). Adverse event rates will be reported as the total number per group, the mean number per patient and the number per category in each group.

### **DMARD escalation protocol**

A step-up DMARD escalation regimen was used for both groups because it matched closely current local practice and fitted best with the 'treat-to-target' ethos of DMARD escalation. In order to avoid being disadvantaged through participation in a clinical study, the DAS28 group were treated using a DMARD escalation approach that has consistently been shown to produce significant clinical benefits (118,146,149,407). Thus, any additional clinical benefits provided by MSUS assessment would have needed to be substantial in order to be measurable at a group level, and significantly different from the benefits experienced by the DAS28 group.

Current UK guidelines restrict the prescription of anti-TNF $\alpha$  blocking therapies to RA patients with persistent high disease activity (DAS28 $>$ 5.1) despite taking 2 non-biologic DMARDs for at least 6 months each (272). In certain circumstances, these guidelines are counter to the concept of aggressively treating until a LDAS or remission target is reached since patients can exhibit clear evidence of active synovitis but may still not qualify for addition of anti-TNF $\alpha$  blocking therapy if the DAS28 score is too low. In order, that the DMARD escalation protocol of this research followed 'tight control' principles the disease activity threshold at which participants in either group could potentially qualify for anti-TNF $\alpha$  blocking therapy was substantially lowered (DAS28 group: DAS28 $<$ 3.2; MSUS: i. PD signal 2 or more joints or ii. 3.2 $<$ DAS28 $\leq$ 5.1 AND SJC $\leq$ 2). However, to ensure that participants had taken two DMARDs for at least 6 months each they were required to convert to subcutaneous methotrexate for a further 3 months if disease activity assessment still exceeded the DMARD escalation threshold despite 3 months of treatment with triple DMARD therapy. Thus, whilst the inclusion of a subcutaneous methotrexate step did ensure study participants had taken 2 non-biologic DMARDs for at least 6 months each before being

considered for anti-TNF $\alpha$  blocking therapy, it may also have delayed commencement of anti-TNF $\alpha$  blocking therapy in the subset of participants who were unlikely to respond. Subsequently, delays in commencement of anti-TNF $\alpha$  blocking therapy may have reduced the studies ability to detect the positive clinical and radiological benefits of anti-TNF $\alpha$  blocking therapy within the remainder of the follow-up period. Most anti-TNF $\alpha$  therapy treated study participants continued to be reviewed until they had completed the 6 months of etanercept therapy they qualified for under the DMARD protocol. However, Sr Stirling, who conducted the outcomes assessment, was blinded to both DMARD treatment regimen and randomisation group so was unable to conduct further disease activity assessments when participants returned for additional reviews that fell outwith the usual 18 month follow-up period.

It is unlikely that the findings of this research will be applicable to all rheumatology services. Some rheumatology centres favour a parallel and step-down DMARD approach (120,143) and in some countries prescribing guidelines allow for the prescription of anti-TNF $\alpha$  blocking therapy at a much earlier stage in the disease process. In both these contexts, MSUS disease activity assessment is very unlikely to alter management since a decision has already been made to commence patients on the highest tier of treatment. In parallel and step-down DMARD regimens, MSUS disease activity assessment may serve an opposite purpose to that proposed by this study; to identify those patients who no longer exhibit evidence of active synovitis in whom it may be possible to reduce DMARD therapy. However, the efficacy of this approach requires to be proven in a randomised trial.

In several countries, after methotrexate failure, the addition of biologic therapy (particularly anti-TNF $\alpha$  blocking therapy) is favoured rather than the addition of sulphasalazine and hydroxychloroquine. Whilst this approach ensures that patients with persistently active disease are offered anti-TNF $\alpha$  blocking therapy at a very early point in their treatment sequence it could potentially lead to unnecessary, and expensive, exposure of patients to the risks of anti-TNF $\alpha$  blocking therapy when they might have responded adequately to triple DMARD therapy instead. Further, this approach would not be feasible under current UK anti-TNF $\alpha$  blocking therapy prescribing restrictions. Following on from the previously described SWEFOT (147,252) and TEAR studies (146), the recent RACAT study has examined whether choosing to add sulphasalazine and hydroxychloroquine instead of etanercept disadvantages RA patients who have experienced an inadequate response to methotrexate monotherapy (408). Overall, there was no significant difference in the change of any of the clinical, functional and radiographic outcome measures after 48 weeks. There was a tendency for the etanercept-methotrexate group to exhibit slightly better mean change in DAS28 after 24 weeks, mean change in HAQ after 48 weeks and higher rates of both DAS28 LDAS and DAS28 remission. Interestingly, therapy switching rates were equal between both groups and patients who switched therapy tended to gain additional clinical response. The RACAT study has carefully reproduced a common clinical scenario (i.e. what is the most appropriate treatment decision after methotrexate monotherapy failure?) in the setting of a randomised, controlled clinical trial. The results support the DMARD escalation regimen chosen for this study, and also support local practice. That is, using triple DMARD therapy following methotrexate monotherapy failure as an interim measure in

methotrexate non-responders should not disadvantage the patient since, in the short term, it does not appear to be associated with worse clinical, functional or radiographic outcomes. Further, the majority of patients (75%) who use triple DMARD therapy following methotrexate monotherapy failure will experience improvements that are similar to those that are experienced by the addition of etanercept but without the additional expense or risk of treatment associated adverse effects.

### **Bias of investigator**

The disease activity assessments and DMARD escalation decisions for all research participants were conducted by a single researcher. Indeed, Dr Dale was the sole rheumatologist that research participants met for the duration of their participation in the research. Consequently, Dr Dale was not blinded to participant's randomisation groups and, through monthly calculation of the DAS28, had insight into how well each participant's RA was responding to treatment. A single investigator should reduce the inherent variability of subjective disease activity assessments and ensure consistency of DMARD escalation decisions. However, there is also a risk that awareness of both group's disease activity levels might introduce an inadvertent bias that favours DMARD escalation in the MSUS group and restricts DMARD escalation in the DAS28 group (especially when the disease activity assessment is on the borderline of the DMARD escalation threshold). Potential bias will have been partly limited by the DMARD escalation protocol which pre-defined DMARD escalation thresholds and the sequence of DMARD escalation. It may have been possible to limit bias further by blinding the Dr Dale to the method of disease activity assessment (i.e. an additional researcher conducts all the MSUS activity assessments) or by having separate clinicians responsible for the assessment and management of each group. However, at the time the research was being devised there was neither the staff, nor the resources, available to make such an arrangement

### **Incomplete DAS28 vs MSUS dataset**

For this thesis, the results pertaining to the agreement between DAS28 and MSUS (Chapter 5) and those pertaining to the impact of MSUS disease activity assessment on ACR core set outcomes (Chapter 6) refer to different datasets. The DAS28 and MSUS comparison was conducted using an incomplete set of data since not all participants had completed the full follow-up period. The available data did provide simultaneous DAS28 and MSUS disease activity assessment data from a substantial number of consultations which, when pooled together, allowed the degree of agreement between both assessment methods to be calculated. Attempts were made to describe the change in MSUS findings over time; however, there was only complete data available for MSUS group participants until follow-up month 15 since a small number of participants (approximately 7, 13% of MSUS group) had not yet attended for later reviews. Thus the presented results for follow-up months 16-18 do not represent a complete dataset and the final findings may be altered (marginally) by the inclusion of all missing data. By comparison, the results presented for the ACR core set comparison were gathered by analysing the complete data set and can reasonably be assumed to be an accurate representation of the response to MSUS guided DMARD escalation. Thus, since the data sets used by the two MSUS comparisons relate to slightly different numbers and sequences of consultations they may not be

directly compared; though, there will be considerable overlap in the implications of the findings from both.

### **Absence of MSUS data for DAS28 group**

Several previous studies have demonstrated that MSUS findings do improve in response to changes in immunomodulatory therapy (220,373,409-413). For this research, MSUS disease activity assessment was only conducted on participants randomised to the MSUS assessment group. It is assumed that the observed changes in MSUS findings of synovial hypertrophy and PD signal relate to the positive influence of escalating doses of DMARD therapy and, by extension, the influence of the preceding disease activity assessment method. However, there is no parallel group against which the changes in MSUS findings can be compared. Furthermore, MSUS assessments were not performed on all patients at regular intervals and therefore the description of the change in MSUS findings over time is based on a patchy and incomplete dataset. In future studies, it would seem appropriate to perform MSUS disease activity assessment on all participants (even those who do not receive MSUS guided DMARD escalation) at every single consultation to: i. determine more clearly how MSUS findings change in response to changes in DMARD therapy and steroid administration; ii. determine whether there is any difference in the rate of change in MSUS findings in the control (non-MSUS assessed) group and iii. determine how often positive clinical responses are associated with persistence of ultrasonographic PD signal.

It is not wholly accurate to report the same MSUS findings as both a disease activity measure and as a separate marker of treatment response since, even though they are clearly linked, they are not independently recorded. For this research, the change in MRI RAMRIS appearances of synovitis and bone marrow oedema between baseline and follow-up month 18 will serve as independent, radiographic measures of treatment response. If a MSUS measure of treatment response were to have been included as an outcome measure, the examination would need to have been performed independently of the disease activity assessment examination and, ideally, over a much more extensive joint set so that the ability of the limited set to represent global disease activity could be determined.

Despite, an incomplete dataset, an irregular frequency of MSUS assessment for some participants, and the preceding concerns relating to the use of MSUS disease activity findings as a surrogate outcome measure, there did seem to be some improvement in MSUS PD findings over time. At the time of analysis, there was only 5 sets of paired 3 and 18 month MSUS data available for analysis; therefore, it was not possible to demonstrate a significant improvement in the MSUS findings over the whole time period. It was possible to demonstrate a significant improvement in PD findings between each participant's first and last MSUS assessment (PD joint count: 1.78 to 1.12, p 0.004); however, this does not represent a constant time period for all participants, and doesn't account for the number of MSUS assessments each participant had undergone. Further, the sensitivity of the MSUS assessment to describe large changes in inflammatory disease activity will have been restricted because the most marked clinical improvements occurred between baseline and follow-up month 3. Comparing changes from

baseline may demonstrate a much larger change in MSUS findings since it also corresponds to a much greater overall change in overt clinical disease activity. MSUS synovial hypertrophy joint count and total score did not show the same degree of improvement as PD findings over the follow-up period. In fact, there was no apparent reduction in either synovial hypertrophy joint count or total score identified. This is contrary to previous longitudinal studies that have suggested synovial hypertrophy findings improve in response to changes in treatment (220,292,332,351,413). There was considerable variation in the size, design, choice of joint set and treatment intervention employed by these studies. Nevertheless, even studies utilising small limited joint sets (such as the previously described US7 count) were able to identify significant reductions in total synovial hypertrophy score in patients undergoing a variety of different treatment changes (including step-up DMARD escalation and commencement of biologic therapy) (292,375).

## **9.2 The Gene Expression Analysis**

It is perhaps odd that the findings of the gene expression analysis are so inconsistent. As an example, the findings suggest that 5 genes were up-regulated in rheumatoid factor positive patients; whereas, there was no evidence of differential expression based upon anti-CCP antibody status, a phenotypic trait which arguably has a much stronger link to RA pathogenesis (398,414). Furthermore, the analysis was unable to recreate previously reported associations between gene expression pattern and particular disease activity states; such as, the identification of a peripheral blood gene expression pattern associated with high disease activity (370,393). The TaSER cohort was considerably larger than many of those in previous studies that have suggested a positive relationship between gene expression and RA phenotype. Furthermore, a very high percentage (96%) of the cohort fulfilled 2010 ACR-EULAR RA classification criteria at first presentation. Therefore, it seems likely that the results represent gene expression profiles from a population of patients who would currently be recognised as suffering from RA. Since the results in relation to disease activity are inconsistent it is possible the multiple testings involved in the statistical analysis techniques will have generated some results by chance, despite prior attempts to normalise the data. Nevertheless, a number of the comparisons did identify suggestions of differential gene expression between comparator groups that were similar to previously reported results (e.g. differential gene expression between gender groups, gene up-regulation in RhF positive participants, up-regulation of LRRN3 (but no other genes) in smokers). Altogether, a number of important additional factors must be considered in relation to these initial findings:

### **Choice of target tissue**

The majority of studies comparing gene expression profile to treatment response have been conducted using either whole peripheral blood (239,355,390,415), PBMCs (234,382-384) from peripheral blood or synovial biopsy tissue (235,236,371). Examining synovial tissue seems sensible since it focuses on the end-target tissue of RA and will not be prone to the same external influences (such as comorbid illnesses and environmental exposures) that may also alter gene expression findings in peripheral blood. Unfortunately, the collection of synovial biopsy samples

continues to require a percutaneous biopsy which many patients may find either unpleasant, and/or unacceptable, and could limit the longitudinal assessment of changes in gene expression over serial synovial tissue samples.

Peripheral blood provides an ideal candidate tissue in which to identify potential new and clinically useful markers since it is relatively easy to obtain and allows repeat testing to track longitudinal changes. Since RA is a systemic disease and activated immune cells are known to migrate between synovial sites and lymphoid organs in peripheral blood and lymph it has been presumed that some features of synovial differential gene expression will also be apparent in peripheral blood. However, the continued circulation of blood also exposes it to processes within every other organ system within the body. Therefore, it is quite likely that any gene expression patterns identified in peripheral blood will represent the collective influence of many different simultaneous physiological and disease-related processes. Clearly, multiple overlapping influences on peripheral blood will increase the potential for variability within the gene expression patterns whilst also reducing its overall specificity. Multiple external influences on peripheral blood gene expression have been identified including physiological factors (such as age, gender, body mass index (358,362), sampling-related factors (such as time of day and relative proportion of blood cell lines(358,362), environmental exposures (such as tobacco smoking, diet (353), and concurrent illness (such as cancer, cardiovascular disease, diabetes mellitus, or infection (416-418). Given the potential for multiple environmental exposures and/or co-morbidities it will be impossible to design a sufficiently large study which accounts for every potential confounder and therefore there may always remain a concern that gene expression patterns identified in peripheral blood are related to a co-existing factor rather than the condition under investigation. Further, two studies, that conducted paired analyses between blood and synovial tissue samples, have demonstrated that evidence of differential gene expression in synovial tissue was not present in peripheral blood gene expression analyses (371,419). Therefore, peripheral blood may not be a suitable tissue for analysis if i. differential gene expression is confined to synovial tissue; ii. the degree of overlap in gene expression patterns between the two tissues is too small to allow peripheral blood gene expression to act as a surrogate for synovial tissue and iii. the specificity of peripheral blood is restricted through exposure to multiple external influences on gene expression.

### **Potential impact of laboratory technique on gene expression**

Inherent variability within the laboratory technique may also contribute to variability in the findings of gene expression experiments. This is important because increased variability in the gene expression patterns will reduce the analyses power to identify true gene expression differences between groups. It is well recognised that differences in technique relating to either laboratory condition, reagent usage or personnel (420,421) contribute to variability in gene expression findings (i.e. batch effect). Crucially, the impact of batch effect on gene expression patterns can produce confounded experiments where it is not possible to distinguish real biological differences in gene expression from variability caused by the experimental technique. Attempts are made to minimise the impact of technical variability on the final gene expression results by standardising the experimental technique: i.e. a small number of laboratory workers, analysing all samples

simultaneously in a single laboratory and using reagents and microarray chips from a single batch. Whilst the laboratory analysis technique can be standardised to minimise variability, additional technical factors, that often cannot be controlled, may also contribute to variability in gene expression findings. Commercially available RNA collection tubes (such as PAXgene RNA or Tempus tubes) stabilise whole blood RNA by lysing cells and preventing further transcript synthesis immediately after blood samples are collected. Indeed, both systems have been shown to deliver RNA transcripts of similar quality, breadth and abundance. However, direct comparison of final gene expression data has consistently shown that there is relatively little overlap in expression patterns identified by either system (422,423). Further, even the use of a single brand of RNA collection tube does not totally prevent variability of gene expression whilst the sample is in transit. When all other factors are equal, the length of period that the PAXgene RNA tubes are kept at room temperature prior to freezing has also been shown to influence final gene expression patterns even though it had no impact upon extracted RNA quality or quantity (424).

The potential impact of non-biological variability (i.e technical factors) needs to be considered in relation to the gene expression analysis outputs of this research in case it had obscured evidence of RA-related differential gene expression. Where possible attempts were made to limit variability in the blood collection and analysis processes by 1. using a single batch of PAXgene RNA collection tubes, 2. using a single batch of Illumina HumanHT-12 v4 Beadchips and 3. performing the RNA extraction, microarray hybridisation and analysis on as few occasions as possible using a standardised workflow. Overall, the blood collection and RNA extraction processes produced RNA of sufficient quality for analysis. Indeed, Agilent Bioanalyser quality control analysis showed that only a single sample returned a RNA integrity number (RIN) less than the acceptable threshold of 7. However, despite these findings and attempts to standardise the experimental technique, a number of additional factors were present that may have influenced the final gene expression outputs:

1. Due to the timing of consultations, and the need to travel between hospital and university laboratory, the period between blood sample collection and final storage was not uniform. The length of time was not specifically recorded though it is estimated that most PAXgene RNA samples remained at room temperature for between 2 and 6 hours
2. Three different members of laboratory staff conducted the RNA extraction and microarray hybridisation; therefore, there was potential for inter-operator variability in laboratory technique
3. Due to the number of samples it was not possible to conduct RNA extraction or hybridisation on a single day; therefore laboratory conditions may not have been identical during the processing of each sample

Previous studies have suggested that gene expression outputs may also be affected by the presence of excess, reticulocyte derived globin (421,423) and the relative proportion of different cell lines within the blood sample (425). Prior to microarray hybridisation, globin reduction techniques may increase the sensitivity of the analysis at the risk of reducing the variability of gene expression. However, for this research, globin reduction was not performed because: 1. it

adds expense to the analysis, 2. it has not been shown to confer any advantages in relation to Illumina Beadchips ([http://www.expressionanalysis.com/images/uploads/tech\\_notes/Illumina\\_Globin\\_Tech\\_Note\\_v2.pdf](http://www.expressionanalysis.com/images/uploads/tech_notes/Illumina_Globin_Tech_Note_v2.pdf)) and 3. risks inducing further variations in gene expression (426).

Considering the complexity of the experimental technique, the multiple different potential sources of biological and technical variability and the inconsistency of current results it seems unlikely that any aspect of whole blood gene expression analysis in its present guise will be adopted into common rheumatological practice. The need to batch samples may prevent the eventual gene expression outputs providing a timely assessment of disease state and the need to carefully standardise laboratory techniques may prove prohibitive in busy clinical laboratories. However, it is unlikely that the necessary laboratory techniques will be optimised for wide spread use until robustly validated gene expression patterns have been identified, that relate to important clinical scenarios and meaningfully influence clinical practice. That being said, gene expression analysis by microarray serve to identify candidate gene profiles for further investigation that are then validated and incorporated within stable and commercially viable analysis techniques (e.g. ELISA)

### **Choice of microarray platform**

Commonly used microarrays have gene probes either fixed to a static plate (e.g. Affymetrix Human Genome U133a Array) or attached to microscopic beads (e.g. Illumina HumanHT Beadchip). Each platform contains a large number of well characterised human genes, recognised variants and gene candidates. However, there is significant variability in the number and breadth of gene probes present on different microarrays. As an illustration, the Affymetrix Human Genome U133a 2.0 Array represents 14,500 human genes, whilst the Illumina HumanHT-12 v2.0 Beadchip represents at least 47,000 genes. Thus, the findings of gene expression experiments can be heavily influenced by the choice of microarray platform. Microarrays with lesser numbers of gene probes risk being insensitive, whereas platforms with very large numbers of probes risk being non-specific. Previous studies have clearly shown that analysing tissue extracts on different microarray platforms on different microarray platforms produces markedly different gene expression patterns (427-429). Therefore, it is difficult to directly compare gene expression findings between experiments if they have been performed on different platforms.

### **Choice of statistical analysis technique**

Studies that describe the relationship between gene expression pattern and RA phenotype have been conducted using a variety of different statistical analysis techniques. Studies with small numbers of patients have tended to use differences in gene fold change to identify differential gene expression, whereas larger studies have tended to use a variety of different techniques based around either t-tests (e.g. significance analysis of microarrays), ANOVA or the Empirical Bayes method (such as Limma). Once again, the choice of statistical analysis technique increases the risk of variability, and therefore reduces the reproducibility, between different experiments. Previous studies that have applied different statistical analysis techniques to the same gene expression data set have shown that there is relatively little overlap (approximately

21%) in the gene lists identified by each technique (430,431). For this research, limma was chosen in preference to other techniques because of its proven performance characteristics. Compared to alternative techniques, limma analysis has been shown to perform consistently for small and large sample sizes, to have a favourable false-positive gene identification rate and to facilitate more complicated analyses between 2 or more comparison groups (431,432). However, clearly, applying an alternative statistical analysis technique may have identified different gene expression findings.

Given the size of the cohort and the number of gene probes on the Illumina HumanHT Beadchip uncontrolled comparisons of expression levels of individual genes between every single participant would lead to a high risk of identifying evidence of differential gene expression by chance (i.e. a false discovery). By using an adjusted p value to correct for multiple comparisons it is possible to control the number of incorrectly identified incidents of differential gene expression, thereby reducing the likelihood of false positive findings and the risk of type I errors (319). Conventionally, an adjusted p value threshold of 5% (0.05) is considered evidence of significant differential gene expression between comparator groups. It is worth noting for this study that a number of the comparisons reported as positive were not achieved until the adjusted p value threshold had been arbitrarily increased to 15% (0.15). Thus increasing the risk that the observed results are not biologically relevant. This particularly relates to the comparisons between baseline expression profile and RhF status and baseline expression profile and 3 month DAS28<3.2. Differences in statistical analysis techniques may be relevant when considering that this research was unable to replicate the findings of previous studies describing the relationship of interferon response gene expression to disease activity in patients treated with rituximab (355,356). The experimental techniques of both studies and this research appear similar and all analysed blood gene expression profiles using the Illumina HumanHT-12 Beadchip (v3 for Vosslamber et al and Raterman et al, v4 for this research). Importantly, whilst this study employed the limma statistical analysis technique, the previous studies both chose the significance analysis of microarray technique instead. The study by Vosslamber et al did identify evidence of differential gene expression using a conventional false discovery threshold of 5%. However, Raterman et al did not identify any evidence of differential gene expression at that threshold and used a 2 fold-difference in expression levels to define differential expression instead. Comparison of gene-fold changes does not take account of variance within the data set and is therefore prone to the influence of outliers and type 1 errors. Clearly the influence of outliers may also be exaggerated in a small cohort, such as the one studied by Raterman et al. However, the proposed interferon gene scores are based upon mean gene expression levels, and not the adjusted p value; therefore, the differences in statistical technique should not have influenced the calculation or application of interferon gene scores.

### **Choice of comparator group descriptors**

This study chose to use DAS28 as the main descriptor of RA global disease activity and treatment response since it is well understood, validated and responsive to changes in disease activity state. Therefore, if this study had identified clear associations between RA phenotype and gene expression profile the potential clinical implications would be readily apparent to everyday

clinicians. However, as previously discussed (Section 1.4), the DAS28 can be relatively insensitive at very low levels of disease activity and also skewed by external, non-inflammatory influences. Therefore, the DAS28 may not have been a specific enough denominator and could have led to grouping inconsistencies. Participants may have been classified as DAS28 remission when in fact they still had active subclinical disease and, similarly, participants may have also been classified as having active disease without evidence of underlying synovitis. In both instances, the disconnect between the presumed and actual disease activity states may also have been reflected in differences of gene expression that biased the overall group gene expression signatures. In fact, the available results suggest that this concern may be unfounded since there was no evidence of differential gene expression between the sub-group of participants with no evidence of synovitis on MSUS and the sub-group with clear, clinically active disease. If there was no significant differences between participants with very evident differences in disease activity phenotype it is unlikely that misclassifying a small number of patients, with closely matched phenotypes, would have significantly skewed the overall gene expression findings for the group.

Previous studies have demonstrated an association between gene expression profile and a number of different methods of classifying both disease activity and treatment response. So far, gene expression profiles in blood and synovial tissue have been correlated to the following disease activity descriptors at a single time point HAQ (377), CRP (371,377), ESR (371) and DAS28 (370,371). Comparisons with treatment response have been more consistent, with previous studies tending to use either change in DAS28 following treatment (234,236,238,355,356,381,382) or EULAR response criteria (241,380,381,415) to form gene expression comparator groups. The most commonly used definition of a good clinical response to treatment has been a fall in DAS28 of at least 1.2 (234,236,238,355,356,381,433) since this is the threshold of the minimally important difference in DAS28. For an individual patient, the gene expression analysis adds little to the perception of disease activity since the DAS28 (which can be calculated during the consultation), has already 'detected' a change. Equally, since current treatment guidelines advocate escalating treatment until patients achieve either low disease activity and/or clinical remission (124,169) accepting the minimally detectable change is inappropriate if patients continue to exhibit moderate (or higher) disease activity and/or clinically active disease. However, it is also worth considering that most studies describing the relationship between gene expression and RA treatment response were conducted in patients with established disease as they commenced biologic therapy whilst recent 'treat-to-target' treatment guidelines relate to the management of the very earliest stages of RA.

Each of the individual components of the DAS28 is subject to variation and external influences and, as discussed previously, may also be associated with specific patterns of gene expression. Thus, whilst clinically useful, it cannot be presumed that the gene expression patterns identified when patients are grouped according to DAS28 are confined to the DAS28 since there are many additional factors that might influence its calculation. Grouping by single variables (such as number or distribution of swollen joints or CRP level) focuses on a single aspect of phenotype and presumes that the expression of this particular phenotype is related to a restricted number of

pathogenetic pathways that, in turn, are related to the activation states of a restricted number of genes. Hence, groupings based upon single variables may increase the chance of identifying between group differences in gene expression. In due course, the gene expression dataset for this study will undergo further interrogation using comparator groups based upon the levels of alternative, single feature, disease activity measures (such as: CRP level, ESR level, number and distribution of swollen joint, number and distribution of tender joints).

## **Validation**

It is acknowledged that the gene expression analysis results presented herein represent an incomplete analysis. Given the nature of the findings - particularly the scarcity of significant positive findings - there has not (yet) been the opportunity to conduct additional analyses that are commonly reported in other studies of gene expression. In particular, there has not been any attempt to validate the presence of differential gene expression by quantitative real-time PCR (qPCR). Equally, there has been no attempt to determine whether any of the genes that did exhibit differential expression between groups is biologically plausible since, so far, none of the gene list data has been entered into either existing pathway analysis software or gene ontology databases. To ensure that no clinically relevant gene signatures are missed, additional analyses are planned whereby comparator groups will be formed based upon single factor disease activity (e.g. CRP) and phenotypic (e.g. tender + swollen joint counts) variables. Equally, since the individual chemical components of the VECTRA-DA relate directly to the recognised pathogenesis of RA, additional gene expression analyses will be performed whereby comparators groups are formed based upon the expression levels of the individual cytokines. Eventually, once the process of identifying clinical scenarios with evidence of differential gene expression has been exhausted, the expression levels of candidate genes will be confirmed by qPCR. Thereafter, the performance of the proposed gene signature will be tested prospectively in an independent validation cohort. If the proposed gene signature maintains the same relationship to clinical phenotype in both training and validation cohorts it is likely that it represents a 'true' finding which, depending upon the clinical context, may be applicable to the routine assessment of RA.

## **9.3 The Multi-Biomarker Disease Activity Test Analysis**

The results of the comparison between the MBDA test, DAS28-ESR and MSUS assessments of global disease activity suggest that the MBDA may provide additional, useful disease activity information in RA patients. It is encouraging that the degree of correlation between MBDA and DAS28-ESR at a single time point, and change over time, is similar to those published by other groups. However, there are several important factors that need should be considered to place the results in context:

### **Temporal relationship of disease activity to DAS28-ESR and MBDA test**

The results for this analysis were collected in parallel to an interventional study that used alternative disease activity measures to guide therapy escalation. Therefore, it is only possible to

describe the correlation between the DAS28-ESR and MBDA and presume that fluctuations in the MBDA test score are in part related to DMARD escalation decisions that were based upon the value of the DAS28-ESR. From these results it is possible to describe the degree of agreement and disagreement between the MBDA test and the corresponding DAS28-ESR. However, given the high prevalence of high MBDA disease activity in participants with low DAS28-ESR disease activity, it is possible that the two measures are not entirely representing the same disease activity states. Further, since the MBDA test was used as an accessory measure of disease activity, it is not possible to comment how the test score will react when DMARD escalation decisions are made on the basis of the MBDA. It is extremely interesting that there is a consistent, significant positive correlation between the MBDA test score and the DAS28-ESR; however, the clinical value (and safety) of the MBDA may not become apparent until its impact on disease outcomes and adverse event rates is tested in a prospective interventional study that uses the MBDA test score to drive DMARD escalation. It is quite possible that targeting low MBDA disease activity leads to a treatment burden or adverse event rate that patients, and clinicians, find unacceptable.

Due to the time points at which research samples were taken it has only been possible to describe the relationship between MBDA score and DAS28-ESR at baseline and after 3 and 18 months of follow-up. Overall, the results suggest that the steepest fall in MBDA occurs between baseline and 3 months of treatment, after which the rate of change is noticeably slower. This may be because the fastest rate of change in disease activity also occurs between baseline and 3 months as patients are established on DMARD therapy and usually also receive frequent doses of systemic and/or intra-articular corticosteroids. Unfortunately, the absence of any intervening MBDA test results between follow-up months 3 and 18 means that it is not possible to discuss in detail how the MBDA fluctuates over time once a stable disease activity state has been attained.

### **Use of DAS28-ESR rather than DAS28-CRP**

The majority of the published studies that have tested the MBDA in clinical practice (383,390,392,394,399) have used the DAS28-CRP as a comparator variable because the CRP is also a component of the MBDA biomarker set. So far, a single study (393) has described the relationship between DAS28-ESR and MBDA test score and demonstrated similar correlations to those described for the DAS28-CRP. For this study the DAS28-ESR was chosen as the comparator variable, since throughout the duration of the clinical study treatment decisions were based upon the DAS28-ESR and these results were more readily available. Overall the correlations between DAS28-ESR and MBDA test score are similar to those of previously published research, so it seems unlikely that the use of DAS28-ESR will have significantly biased the findings. Further, the results suggest that the relationship between MBDA and disease activity persists when the interdependence on the C-reactive protein level was removed.

### **Sensitivity of MBDA in relation to alternative disease activity measures**

As previously discussed, the findings of this study have been consistent with other previous research; that is, during DAS28-ESR LDAS and remission the MBDA consistently detects a

higher level of disease activity in a subset of patients. In addition, this study also found that the MBDA continues to report moderate-to-high disease activity in a significant proportion of participants who did not have MSUS evidence of active synovitis. This suggests that the MBDA may be more sensitive even than MSUS for the detection of on-going synovitis. However, it should be noted that there was only a relatively small number of occasions (n = 35) when there was simultaneous DAS28-ESR and MBDA data available for comparison and the finding should be tested in a larger cohort. Furthermore, once the 18 month MRI outcomes become available, it will be extremely interesting to determine whether the incidence of MBDA moderate-high disease activity continues to exceed the incidence of on-going synovitis detected in corresponding MRI images. Equally, if MBDA does demonstrate the presence of active synovitis the presumption is that there should be a greater degree of agreement between MRI and MBDA findings than presently demonstrated between DAS28-ESR and MSUS. MRI may prove a more appropriate validation measure against which to test the MBDA since it is the most sensitive imaging technique currently available for the identification of synovitis and may help differentiate between how many of the elevated MBDA scores are false positives and how many of the DAS28 LDAS scores are in fact false negatives

### **MSUS joint set**

All of the comparisons between MSUS and MBDA disease activity assessments have been conducted using the findings from the limited MSUS joint set of the main clinical study. This relies heavily upon the presumption that the proposed 14 joint set provides an accurate representation of global disease activity. Certainly, the results of the MSUS analysis demonstrate that the proposed joint set was a more sensitive measure of global disease activity than the corresponding DAS28-ESR. However, the ability of the joint set to fully represent global ultrasonographic activity has not been specifically tested against a much more extensive joint set. The degree of disagreement between the MBDA and MSUS assessments is strikingly high (especially in participants with high-moderate MBDA disease activity) and does question how much of the MBDA elevation is due to the presence of synovitis. However, the degree of disagreement can only be considered in relation to MBDA and the joint set proposed by this research; it can not relate to all methods of MSUS disease activity assessment. An alternative explanation for the disagreement between MBDA and MSUS assessments may be that the MSUS assessment was limited to too few joints and that in many cases of inactive MSUS disease the elevation of the MBDA was related to active synovitis within joints that were not included by the MSUS examination. The available results provide a useful initial assessment of the degree of agreement between MBDA and imaging evidence of active synovitis. A more accurate assessment may have been gained by using a much more extensive joint set to give a much clearer view of overall inflammatory joint disease activity rather than trying to extrapolate from an untested, limited set.

## **10. Conclusions**

All of the results presented within this thesis are based upon one small population of patients with newly diagnosed rheumatoid arthritis. The initial intention of the TaSER Study had been to test the impact of musculoskeletal ultrasound disease activity assessment on DMARD escalation decisions and short-medium term outcomes in a randomised clinical trial setting. The systematic collection of clinical, functional and radiological outcome data has allowed a very detailed description of how each randomisation group has responded to the DMARD therapy administered over the course of the follow-up period. The routine collection and storage of a variety of additional blood samples at set points throughout the follow-up period has provided an opportunity to also compare how certain blood profiles respond to DMARD therapy at important decision points during a participant's treatment sequence (i.e. after the first 3 months of DMARD monotherapy and after 18 months of step-up DMARD therapy). Further, the coexistence of corresponding clinical data has provided an opportunity to integrate the various clinical, ultrasonographic, gene expression and MBDA data sets to determine whether particular clinical and ultrasonographic features of RA phenotype (particularly disease activity state) are also associated with specific patterns of expression within blood. Whilst the list of analyses has not been exhaustive it has led to the scope of the research evolving to consider whether novel blood analysis methods might provide additional means of assessing RA phenotype and/or disease activity.

Based upon the available results it is possible to make the following conclusions in the following subject areas:

#### **Study cohort**

- The cohort recruited to the study were recognisable as a modern early RA population, having typical presenting clinical features and a very high percentage of participants who fulfilled ACR-EULAR 2010 classification criteria.
- Compared to previous early arthritis strategic treatment studies from the same area, the TaSER cohort reported generally shorter disease durations and exhibited lesser levels of disease activity

#### **The impact of MSUS assessment on workload**

- The incorporation of MSUS disease activity assessment into routine rheumatological practice has the potential to significantly impact upon clinical workload by lengthening assessment time and altering clinical work patterns
- Patients with high disease activity, or moderate disease activity and clinically evident synovitis, require relatively few MSUS assessments. The majority (89%) of MSUS assessments are performed for disease monitoring purposes in patients with DAS28 LDAS and a small minority (11%) are performed in patients with moderate disease activity but minimal clinical evidence of synovitis

### **Agreement between MSUS and DAS28ESR and impact on treatment decisions**

- DAS28 and MSUS disease activity assessments agreed on the disease state, and lead to the same treatment decision, on 71% of occasions. In the remaining 29% of occasions, MSUS disease activity findings suggested a disease activity state, and supported a treatment decision, that was contrary to that suggested by the DAS28
- In DAS28 LDAS and remission states, approximately one quarter of MSUS disease activity assessments identified evidence of active disease and therefore supported DMARD escalation. Conversely, in instances of DAS28 moderate disease activity but minimal clinical synovitis, approximately two-thirds of MSUS assessments did not identify evidence of active disease. Therefore, in DAS28 LDAS states regular MSUS disease activity assessment will lead to patients receiving a higher burden of immunomodulatory therapy and, in DAS28 moderate disease activity states without clinical synovitis regular MSUS disease activity assessment will lead to patients receiving a lesser burden of immunomodulatory therapy
- A significant proportion of participants in both DAS28 LDAS and moderate disease activity will exhibit persistently negative MSUS disease activity assessments and could be considered as being in ultrasonographic low disease activity, if not remission
- The radiocarpal and index MCP joints were most likely to exhibit evidence of PD signal

### **Change in MSUS findings over time**

- Power Doppler joint counts and total PD scores continue to fall from follow-up month 3 onwards.
- The degree of improvement of synovial hypertrophy joint count and total score is much less marked than the degree of improvement of PD joint count and total PD score over the same time period

### **Impact of MSUS assessment on clinical outcomes**

- Clinical and laboratory measures of disease activity and functional ability improved significantly in both groups. However, compared to DAS28 steered DMARD escalation, MSUS steered therapy was not associated with significantly better disease activity assessments at any time point nor a statistically greater mean improvement in DAS44 over the whole of the follow-up period. There were no statistically significant between group differences in DAS44, 44 swollen joint count, Ritchie articular index, ESR, CRP, pain 10cm VAS nor global health 10cm VAS at any of the time points. At 18 months follow-up, the MSUS group reported marginally, but not statistically significant, lower HAQ scores and had experienced a mean change in HAQ that was greater than the minimally important difference
- The proportion of participants achieving DAS44 LDAS was similar for both assessment groups at each time point. After 18 months follow-up, MSUS steered DMARD therapy was

associated with a significantly higher rate of DAS44 remission than DAS28 steered therapy. The MSUS group exhibited a marginally higher rate of EULAR good responses and were marginally more likely to have exhibited DAS44 remission on 5 or 6 separate, and/or consecutive, occasions

- On the basis of the clinical outcome results, it is not possible to support the routine use of MSUS disease activity assessment in regular practice since it is has not been associated with significantly better short-medium term outcomes. MSUS disease activity assessment may continue to have an important role in the subset of RA patients where there remains doubt about the clinical assessment of disease activity
- The impact of MSUS driven DMARD therapy on adverse event rates, the individual components of the MRI RAMRIS score and the modified Sharp Score remains to be determined. The preceding statement about the routine use of MSUS disease activity assessment may be revised if MSUS steered DMARD therapy is shown to be associated with either lesser progression of damage on MRI RAMRIS and plain x-rays and/or greater improvements in RAMRIS measures of synovitis and bone marrow oedema.

#### **Relationship between whole blood gene expression profile and baseline phenotype**

- Gender groupings were associated with the most evidence of differential gene expression. Baseline rheumatoid factor status and smoking status was associated with differential expression of small numbers of genes
- Baseline anti-CCP antibody status and plain radiology erosion status were not associated with evidence of differential gene expression

#### **Relationship between whole blood gene expression and disease activity**

- Extremes of DAS28 defined disease activity were not associated with evidence of differential gene expression at any of the time points
- There was no evidence of differential gene expression between groups of patients with clinically active disease and either DAS28 remission or MSUS remission. Therefore, whole blood gene expression profiles do not appear to differentiate between active and inactive disease activity states
- DMARD monotherapy was associated with dynamic changes in gene expression profile between baseline and month 3.
- Continuing to exhibit a DAS28 of 3.2 or higher after 3 months of DMARD monotherapy was associated with up-regulation of 3 genes in baseline samples. There were no other predictive associations identified between gene expression profile and subsequent clinical course

- The baseline interferon response gene signature was not predictive of 3 or 12 month DAS28 response to DMARD therapy. Furthermore, there was no relationship between the change in the score and the corresponding change in DAS28 between 3, 6 or 12 months. These findings are contrary to previously reported findings in relation to rituximab therapy
- It is not currently possible to conclude that whole blood gene expression profile has a consistent relationship with either RA phenotype, disease activity or clinical course. Therefore, at present, it is unlikely that whole blood gene expression profiling could be used in clinical practice to provide additional, useful prognostic or disease activity information. Nevertheless, it is likely that whole blood gene expression profiling may continue to be used as an investigative technique in the attempt to better understand RA pathogenesis and the systemic manifestations of the disease

#### **Relationship of MBDA to DAS28-ESR and MSUS disease activity**

- The MBDA score improved significantly following commencement of step-up DMARD therapy and in parallel to the DAS28-ESR. Commencement of etanercept coincided with further improvements in the MBDA score
- The MBDA score correlates positively with the corresponding DAS28-ESR at a single time point. Furthermore, the change in the MBDA test score also correlates positively with the change in the DAS28-ESR score over the same time period. These findings were similar to previously published results comparing the MBDA score and the DAS28-CRP
- The MBDA test may be more sensitive than DAS28-ESR and MSUS assessment for the detection of on-going synovitis. The relationship of the MBDA test score to MRI measures of synovitis and bone marrow oedema remains to be determined
- Baseline MBDA test score may be an independent predictor of short-medium term clinical response in RA patients commencing DMARD

## Areas for Future Work

The results generated by this research have gone some way to addressing several important clinical questions in relation to the assessment and management of early rheumatoid arthritis. However, there inevitably remain additional areas that require further investigation and consideration. Following on from the completion of this thesis, it is my intention to continue investigating the following areas:

### Study Cohort

- Statistical comparison of baseline presenting features between the TICORA, TEAR and TaSER cohorts to describe how the characteristics of patients recruited to clinical studies of early RA in a single centre have changed over 14 years

### The impact of MSUS assessment on outcomes

- Additional review of all participants to the TaSER study 5 years after recruitment to determine if there is divergence in clinical and radiological outcomes over medium-long term follow-up. Data to be collected includes – clinical disease activity assessment (DAS44), functional and quality of life assessments (HAQ and EQ-5D), estimation of work ability, plain radiographs of hands and feet (for grading by modified Sharp Score), assessment of co-morbidity (including serious illness) and mortality, incidence of orthopaedic surgery and assessment of DMARD / biologic therapy burden. If there is evidence of divergence of outcomes at 5 years, further follow-up will be conducted at 10 years.
- Completion of grading of radiological outcomes: baseline and 18 month plain radiographs to be graded according to modified Sharp Score and baseline and 18 month MRI images to be graded according to OMERACT RAMRIS system. Additional comparisons to include correlation between baseline RAMRIS synovitis and bone marrow oedema score and subsequent change in RAMRIS synovitis score and total modified Sharp score.

### Relationship between whole blood gene expression and RA phenotype

- Subgroup comparisons restricted to participants with particular phenotypic descriptors to determine whether correcting for potential confounders of gene expression increases the likelihood of identifying evidence of differential gene expression. Potential comparisons to include:
  - limited to anti-CCP antibody positive participants, grouped by DAS28-ESR disease activity at a single time point (baseline and/or 3 months) with between group comparisons of gene expression conducted between high/low and upper/lower quartile DAS28 groups
  - limited to RhF positive participants, grouped by DAS28-ESR disease activity at a single time point (baseline and/or 3 months) with between group comparisons of gene expression conducted between high/low and upper/lower quartile DAS28 groups

- limited to participants initially treated with methotrexate, grouped by DAS28-ESR disease activity at a single time point (baseline and/or 3 months) with between group comparisons of gene expression conducted between high/low and upper/lower quartile DAS28 groups
- Gathering of additional peripheral blood biomarker datasets, to include:
  - DNA genotyping (potential for targeted ImmunoChip analysis) – baseline samples
  - serum metabolomics – baseline, 3 month and 18 month samples
  - fine-specificity anti-CCP isotyping (verbal agreement to collaborate with Dr J Sokolove, University of Stanford) – baseline samples

The resulting outputs of these analyses will allow a very detailed description of the genetic and metabolic layers that are measurable in the peripheral blood of patients with early RA. The relationship of each dataset will be compared to RA baseline phenotype, disease activity at a single time point and clinical course to determine whether any of the datasets provide consistent additional information regarding phenotype or disease state. Eventually, once all datasets are available an integrative analysis across all baseline layers of the dataset will be conducted to determine whether panels of biomarkers associate with short-medium term treatment outcome. The integrative analysis will be conducted with assistance from the Biostatisticians of the Glasgow Polyomics Facility

#### **Relationship of MBDA to RA disease activity**

- Comparison between 18 month MRI RAMRIS findings and 18 month MBDA score to determine degree of agreement. Particular focus on the relationship between high and moderate MBDA scores and MRI RAMRIS synovitis score
- Additional future areas of interest could include:
  - gene expression profile comparison between participants with different levels of MBDA disease activity
  - interventional clinical trial in early RA to determine if treating to MBDA low disease activity is associated with improved clinical and radiological outcomes. Participants would be randomised to either MBDA or DAS28-CRP disease activity assessment groups and would all receive a step-up DMARD-biologic escalating regimen aiming for either low MBDA or DAS28-CRP disease activity

It has been an enormous, and highly enjoyable, privilege to spend time conducting the TaSER Study. Hopefully, the results described herein will provide a meaningful contribution to the on-going debate on how best to assess, monitor and treat patients with early rheumatoid arthritis. Taking part in this research has provided a vast number of additional experiences which I hope will remain with me throughout the rest of my clinical career. I have received training in the design, conduct and analysis of clinical trials and have also developed a degree of understanding of the use and limitations of musculoskeletal ultrasound. I have received extremely valuable

training in how best to manage patients with new diagnoses of rheumatoid arthritis: including - the use of aggressive step-up DMARD regimens, the inherent problems and concerns relating to combination DMARD therapy, the management of intercurrent illness and adverse events and how best to counsel and manage the expectations of patients during their treatment course. Throughout this I have also developed a better understanding of the natural course of rheumatoid arthritis and its impact on patient's daily lives. Ultimately, I hope to apply all of these experiences to improve the care that I offer all patients with rheumatoid arthritis throughout the rest of my career. I also hope to build upon my initial research experience and remain as a research-active clinician with a focus on performing clinically focussed clinical trials that address pertinent issues relating to the disease course and management of inflammatory arthritides.

## References

1. Ronnelid J. Longitudinal analysis of citrullinated protein/peptide antibodies (anti-CP) during 5 year follow up in early rheumatoid arthritis: anti-CP status predicts worse disease activity and greater radiological progression. *Annals of the Rheumatic Diseases*. 2005 Dec 1;64(12):1744–9.
2. Silman A, Pearson J. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res*. 2002 May 9;4((suppl 3)):S265–72.
3. van der Heijde D. Erosions versus joint space narrowing in rheumatoid arthritis: what do we know? *Annals of the Rheumatic Diseases*. 2011 Feb 21;70(Suppl 1):i116–8.
4. Lee D, Weinblatt M. Rheumatoid arthritis. *The Lancet*. 2001 Sep 15;358:903–11.
5. Smolen JS, van der Heijde DM, ALETAHA D, Xu S, Han J, Baker D, et al. Progression of radiographic joint damage in rheumatoid arthritis: independence of erosions and joint space narrowing. *Annals of the Rheumatic Diseases*. 2009 Sep 11;68(10):1535–40.
6. Hani El-Gabalawy MD FPOMAI. The preclinical stages of RA: lessons from human studies and animal models. *Best Pract Res Clin Rheumatol*. Elsevier Ltd; 2009 Feb 1;23(1):49–58.
7. Nielen MMJ, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MHMT, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: A study of serial measurements in blood donors. *Arthritis Rheum*. 2004;50(2):380–6.
8. Gorman JD, Lum RF, Chen JJ, Suarez-Almazor ME, Thomson G, Criswell LA. Impact of shared epitope genotype and ethnicity on erosive disease: A meta-analysis of 3,240 rheumatoid arthritis patients. *Arthritis Rheum*. 2004;50(2):400–12.
9. Shmerling R, Delbanco T. The rheumatoid factor: an analysis of clinical utility. *The American Journal of Medicine*. 1991 Nov 7;91:528–34.
10. Schellekens G, Visser H, de Jong B, van den Hoogen F, Hazes J, Breedveld F, et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum*. 2000 Jan 1;43(1):155–63.
11. Arnett F, Edworthy S, Bloch D, Fries J, Cooper N, Healey L, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum*. 1987 Mar 3;31(3):315–24.
12. Machold KP, Stamm TA, Eberl GJM, Nell VKP, Dunky A, Uffmann M, et al. Very recent onset arthritis—clinical, laboratory, and radiological findings during the first year of disease. *The Journal of Rheumatology*. 2002 Nov;29(11):2278–87.
13. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 Rheumatoid arthritis classification criteria: An American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum*. 2010 Aug 10;62(9):2569–81.
14. Cader MZ, Filer A, Hazlehurst J, de Pablo P, Buckley CD, Raza K. Performance of the 2010 ACR/EULAR criteria for rheumatoid arthritis: comparison with 1987 ACR criteria in a very early synovitis cohort. *Annals of the Rheumatic Diseases*. 2011 May 2;70(6):949–55.
15. Visser H, le Cessie S, Vos K, Breedveld FC, Hazes JMW. How to diagnose rheumatoid arthritis early: A prediction model for persistent (erosive) arthritis. *Arthritis Rheum*. 2002 Feb 5;46(2):357–65.

16. van der Helm-vanMil AHM, le Cessie S, van Dongen H, Breedveld FC, Toes REM, Huizinga TWJ. A prediction rule for disease outcome in patients with Recent-onset undifferentiated arthritis: How to guide individual treatment decisions. *Arthritis Rheum.* 2007;56(2):433–40.
17. van der Helm-van Mil AHM, Detert J, Cessie SL, Filer A, Bastian H, Burmester GR, et al. Validation of a prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: Moving toward individualized treatment decision-making. *Arthritis Rheum.* 2008 Aug;58(8):2241–7.
18. van Aken J. Comparison of long term outcome of patients with rheumatoid arthritis presenting with undifferentiated arthritis or with rheumatoid arthritis: an observational cohort study. *Annals of the Rheumatic Diseases.* 2006 Jan 1;65(1):20–5.
19. Verpoort KN, van Dongen H, Allaart CF, Toes REM, Breedveld FC, Huizinga TWJ. Undifferentiated arthritis--disease course assessed in several inception cohorts. *Clin. Exp. Rheumatol.* 2004 Sep;22(5 Suppl 35):S12–7.
20. Scott DL, Symmons DP, Coulton BL, Popert AJ. Long-term outcome of treating rheumatoid arthritis: results after 20 years. 1987. pp. 1108–11.
21. Rojas-Serrano J, Pérez LL, García CG, Moctezuma F, Álvarez-Hernández E, Vázquez-Mellado J, et al. Current smoking status is associated to a non-ACR 50 response in early rheumatoid arthritis. A cohort study. *Clin. Rheumatol.* 2011 Dec;30(12):1589–93.
22. Saevarsdottir S, Wedrén S, Seddighzadeh M, Bengtsson C, Wesley A, Lindblad S, et al. Patients with early rheumatoid arthritis who smoke are less likely to respond to treatment with methotrexate and tumor necrosis factor inhibitors: Observations from the Epidemiological Investigation of Rheumatoid Arthritis and the Swedish Rheumatology Reg. *Arthritis Rheum.* 2010 Dec 28;63(1):26–36.
23. Drossaers-Bakker KW, de Buck M, van Zeben D, Zwinderman AH, Breedveld F, Hazes JMW. Long-term course and outcome of functional capacity in rheumatoid arthritis. *Arthritis Rheum.* 1999 Sep 9;42(9):1854–60.
24. Welsing P, van Gestel A, Swinkels H, Kiemeny L, van Riel P. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. *Arthritis Rheum.* 2001 Sep 9;44(9):2009–17.
25. Wick MC. Relationship between inflammation and joint destruction in early rheumatoid arthritis: a mathematical description. *Annals of the Rheumatic Diseases.* 2004 Jul 1;63(7):848–52.
26. Finckh A, Liang MH, van Herckenrode CM, de Pablo P. Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: A meta-analysis. *Arthritis Rheum.* 2006;55(6):864–72.
27. Fuchs H, Kaye J, Callahan L, Nance E, Pincus T. Evidence of significant radiographic damage in rheumatoid arthritis within the first 2 years of disease. *Journal of Rheumatology.* 1989. pp. 585–91.
28. van der Heijde D, van Leeuwen M, van Riel P, Koster A, van't Hof M, van Ruswijk M, et al. Biannual radiographic assessments of hands and feet in a three year prospective follow-up of patients with early rheumatoid arthritis. *Arthritis Rheum.* 1992 Jan 1;36(1):26–34.
29. Gossec L. Prognostic factors for remission in early rheumatoid arthritis: a multiparameter prospective study. *Annals of the Rheumatic Diseases.* 2004 Mar 17;63(6):675–80.
30. McGonagle D, Conaghan P, O'Connor P, Gibbon W, Green M, Wakefield R, et al.

The relationship between synovitis and bone changes in early untreated rheumatoid arthritis. *Arthritis Rheum.* 1999 Aug 1;42(8):1706–11.

31. Combe B, Cantagrel A, Goupille P, Bozonnat M-C, Sibilia J, Eliaou J-F, et al. Predictive factors of 5-year health assessment questionnaire disability in early rheumatoid arthritis. *The Journal of Rheumatology.* 2003 Nov;30(11):2344–9.
32. Verstappen SMM, Symmons DPM. What is the outcome of RA in 2011 and can we predict it? *Best Pract Res Clin Rheumatol.* 2011 Aug;25(4):485–96.
33. Amos RS, Constable TJ, Crockson RA, Crockson AP, McConkey B. Rheumatoid arthritis: relation of serum C-reactive protein and erythrocyte sedimentation rates to radiographic changes. *Br Med J.* 1977 Jan 22;1(6055):195–7.
34. Symmons D, Barrett E, Bankhead C, Scott D, Silman A. The incidence of rheumatoid arthritis in the united kingdom: results from the Norfolk Arthritis Register. *Br. J. Rheumatol.* 1994 Aug 1;33(8):735–9.
35. Plant MJ, Williams AL, O'Sullivan MM, Lewis PA, Coles EC, Jessop JD. Relationship between time-integrated C-reactive protein levels and radiologic progression in patients with rheumatoid arthritis. *Arthritis Rheum.* 2000 Jul;43(7):1473–7.
36. Sokka T, Toloza S, Cutolo M, Kautiainen H, Makinen H, Gogus F, et al. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. *Arthritis Res Ther.* 2009;11(1):R7.
37. Conaghan PG, Hensor EMA, Keenan AM, Morgan AW, Emery P, the YEAR Consortium. Persistently moderate DAS-28 is not benign: loss of function occurs in early RA despite step-up DMARD therapy. *Rheumatology.* 2010 Sep 10;49(10):1894–9.
38. Hallert E, Thyberg I, Hass U, Skargen E, Skogh T. Comparison between women and men with recent onset rheumatoid arthritis of disease activity and functional ability over two years (the TIRA project). *Annals of the Rheumatic Diseases.* 2003 Jul 1;62(7):667–70.
39. IIKUNI N, SATO E, HOSHI M, INOUE E, TANIGUCHI A, HARA M, et al. The Influence of Sex on Patients with Rheumatoid Arthritis in a Large Observational Cohort. *The Journal of Rheumatology.* 2009 Mar 13;36(3):508–11.
40. Jawaheer D, MARANIAN P, PARK G, LAHIFF M, AMJADI SS, Paulus HE. Disease Progression and Treatment Responses in a Prospective DMARD-naive Seropositive Early Rheumatoid Arthritis Cohort: Does Gender Matter? *The Journal of Rheumatology.* 2010 Dec 1;37(12):2475–85.
41. Salaffi F, Carotti M, Ciapetti A, Gasparini S, Filippucci E, Grassi W. Relationship between time-integrated disease activity estimated by DAS28-CRP and radiographic progression of anatomical damage in patients with early rheumatoid arthritis. *BMC Musculoskelet Disord.* 2011;12:120.
42. Naz SM, Farragher TM, Bunn DK, Symmons DPM, Bruce IN. The influence of age at symptom onset and length of followup on mortality in patients with recent-onset inflammatory polyarthritis. *Arthritis Rheum.* 2008 Mar 27;58(4):985–9.
43. Pease CT, Bhakta BB, Devlin J, Emery P. Does the age of onset of rheumatoid arthritis influence phenotype?: a prospective study of outcome and prognostic factors. *Rheumatology (Oxford).* 1999 Mar;38(3):228–34.
44. Kuiper S, van Gestel AM, Swinkels HL, de Boo TM, da Silva JA, van Riel PL. Influence of sex, age, and menopausal state on the course of early rheumatoid arthritis. *The Journal of Rheumatology.* 2001 Aug;28(8):1809–16.

45. Camacho EM, Verstappen SMM, Lunt M, Bunn DK, Symmons DPM. Influence of age and sex on functional outcome over time in a cohort of patients with recent-onset inflammatory polyarthritis: Results from the Norfolk arthritis register. *Arthritis Care Res.* 2011 Nov 29;63(12):1745–52.
46. Paimela L, Palosuo T, Aho K, Lukka M, Kurki P, Leirisalo-Repo M, et al. Association of autoantibodies to filaggrin with an active disease in early rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 2001 Jan;60(1):32–5.
47. Combe B, Dougados M, Goupille P, Cantagrel A, Eliaou JF, Sibilia J, et al. Prognostic factors for radiographic damage in early rheumatoid arthritis: a multiparameter prospective study. *Arthritis Rheum.* 2001 Aug;44(8):1736–43.
48. Masi AT, Maldonado-Cocco JA, Kaplan SB, Feigenbaum SL, Chandler RW. Prospective study of the early course of rheumatoid arthritis in young adults: comparison of patients with and without rheumatoid factor positivity at entry and identification of variables correlating with outcome. *Seminars in Arthritis and Rheumatism.* 1976 May;4(4):299–326.
49. Innala L, Kokkonen H, Eriksson C, Jidell E, Berglin E, Dahlqvist SR. Antibodies against mutated citrullinated vimentin are a better predictor of disease activity at 24 months in early rheumatoid arthritis than antibodies against cyclic citrullinated peptides. *The Journal of Rheumatology.* 2008 Jun;35(6):1002–8.
50. Erhardt C, Mumford P, Venables P, Maini R. Factors predicting a poor life prognosis in rheumatoid arthritis: an eight year prospective study. *Annals of the Rheumatic Diseases.* 1989 Jan 1;48:7–13.
51. Lindqvist E. Prognostic laboratory markers of joint damage in rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 2005 Feb 1;64(2):196–201.
52. van der Helm-van Mil AHM, Verpoort KN, Breedveld FC, Toes REM, Huizinga TWJ. Antibodies to citrullinated proteins and differences in clinical progression of rheumatoid arthritis. *Arthritis Res Ther.* 2005;7(5):R949–58.
53. Bukhari M, Thomson W, Naseem H, Bunn D, Silman A, Symmons D, et al. The performance of anti-cyclic citrullinated peptide antibodies in predicting the severity of radiologic damage in inflammatory polyarthritis: Results from the Norfolk Arthritis Register. *Arthritis Rheum.* 2007;56(9):2929–35.
54. Lard LR, van Gaalen FA, Schonkeren JJM, Pieterman EJ, Stoeken G, Vos K, et al. Association of the -2849 interleukin-10 promoter polymorphism with autoantibody production and joint destruction in rheumatoid arthritis. *Arthritis Rheum.* 2003 Jul 2;48(7):1841–8.
55. Wessels JAM, Kooloos WM, Jonge RD, De Vries-Bouwstra JK, Allaart CF, Linssen A, et al. Relationship between genetic variants in the adenosine pathway and outcome of methotrexate treatment in patients with recent-onset rheumatoid arthritis. *Arthritis Rheum.* 2006;54(9):2830–9.
56. Potter C, Cordell HJ, Barton A, Daly AK, Hyrich KL, Mann DA, et al. Association between anti-tumour necrosis factor treatment response and genetic variants within the TLR and NF B signalling pathways. *Annals of the Rheumatic Diseases.* 2010 Jun 21;69(7):1315–20.
57. Tan RJJ, Gibbons LJ, Potter C, Hyrich KL, Morgan AW, Wilson AG, et al. Investigation of rheumatoid arthritis susceptibility genes identifies association of AFF3 and CD226 variants with response to anti-tumour necrosis factor treatment. *Annals of the Rheumatic Diseases.* 2010 May 24;69(6):1029–35.
58. Karlson EW, Lee IM, Cook NR, Manson JE, Buring JE, Hennekens CH. A retrospective cohort study of cigarette smoking and risk of rheumatoid arthritis in female health professionals. *Arthritis Rheum.* 1999 May;42(5):910–7.

59. Ballara S, Taylor PC, Reusch P, Marmé D, Feldmann M, Maini RN, et al. Raised serum vascular endothelial growth factor levels are associated with destructive change in inflammatory arthritis. *Arthritis Rheum.* 2001 Sep;44(9):2055–64.
60. Pedersen M, Jacobsen S, Garred P, Madsen HO, Klarlund M, Svejgaard A, et al. Strong combined gene–environment effects in anti–cyclic citrullinated peptide–positive rheumatoid arthritis: A nationwide case–control study in Denmark. *Arthritis Rheum.* 2007;56(5):1446–53.
61. Masdottir B. Smoking, rheumatoid factor isotypes and severity of rheumatoid arthritis. *Rheumatology.* 2000 Nov 1;39(11):1202–5.
62. Mikuls TR, Hughes LB, Westfall AO, Holers VM, Parrish L, van der Heijde D, et al. Cigarette smoking, disease severity and autoantibody expression in African Americans with recent-onset rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 2008 Nov;67(11):1529–34.
63. Dale J, Porter D. Pharmacotherapy: concepts of pathogenesis and emerging treatments. Optimising the strategy of care in early rheumatoid arthritis. *Best Pract Res Clin Rheumatol.* 2010 Aug;24(4):443–55.
64. Anderson JJ, Wells G, Verhoeven AC, Felson DT. Factors predicting response to treatment in rheumatoid arthritis: the importance of disease duration. *Arthritis Rheum.* 2000 Jan;43(1):22–9.
65. M tt nen T, Hannonen P, Korpela M, Nissil M, Kautiainen H, Ilonen J, et al. Delay to institution of therapy and induction of remission using single-drug or combination-disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. *Arthritis Rheum.* 2002 Apr 5;46(4):894–8.
66. Green M, Marzo-Ortega H, McGonagle D, Wakefield R, Proudman S, Conaghan P, et al. Persistence of mild, early inflammatory arthritis: the importance of disease duration, rheumatoid factor, and the shared epitope. *Arthritis Rheum.* 1999 Oct;42(10):2184–8.
67. Munro R, Hampson R, McEntegart A, Thomson EA, Madhok R, Capell H. Improved functional outcome in patients with early rheumatoid arthritis treated with intramuscular gold: results of a five year prospective study. *Annals of the Rheumatic Diseases.* 1998 Feb 1;57(2):88–93.
68. van Leeuwen MA, van Rijswijk MH, van der Heijde DM, Meerman Te GJ, van Riel PL, Houtman PM, et al. The acute-phase response in relation to radiographic progression in early rheumatoid arthritis: a prospective study during the first three years of the disease. *Br. J. Rheumatol.* 1993 Jun;32 Suppl 3:9–13.
69. Emery P, Dorner T. Optimising treatment in rheumatoid arthritis: a review of potential biological markers of response. *Annals of the Rheumatic Diseases.* 2011 Nov 2;70(12):2063–70.
70. Taylor P. Investigational biologic markers in the diagnosis and assessment of rheumatoid arthritis. O'dell J, editor. *Uptodate.com.* 2010 Sep 21;:1–14.
71. Bang H, Egerer K, Gauliard A, Lüthke K, Rudolph PE, Fredenhagen G, et al. Mutation and citrullination modifies vimentin to a novel autoantigen for rheumatoid arthritis. *Arthritis Rheum.* 2007;56(8):2503–11.
72. Rook G, Steele J, Brealey R, Whyte A, Isenberg D, Sumar N, et al. Changes in IgG glycoform levels are associated with remission of arthritis during pregnancy. *Journal of Autoimmunity.* 1991 Oct 1;4(5):779–94.
73. Constantin A, Lauwers-Canc s VR, Navaux FDR, Abbal M, Meerwijk JV, Mazi res B, et al. Stromelysin 1 (matrix metalloproteinase 3) and HLA-DRB1 gene polymorphisms: Association with severity and progression of rheumatoid arthritis in

- a prospective study. *Arthritis Rheum.* 2002 Jul 11;46(7):1754–62.
74. van Leeuwen MA, Westra J, Limburg PC, van Riel PL, van Rijswijk MH. Clinical significance of interleukin-6 measurement in early rheumatoid arthritis: relation with laboratory and clinical variables and radiological progression in a three year prospective study. *Annals of the Rheumatic Diseases.* 1995 Aug;54(8):674–7.
  75. Paimela L, Heiskanen A, Kurki P, Helve T, Leirisalo-Repo M. Serum hyaluronate level as a predictor of radiologic progression in early rheumatoid arthritis. *Arthritis Rheum.* 1991 Jul;34(7):815–21.
  76. Green MJ. Serum MMP-3 and MMP-1 and progression of joint damage in early rheumatoid arthritis. *Rheumatology.* 2003 Jan 1;42(1):83–8.
  77. Månsson B, Carey D, Alini M, Ionescu M, Rosenberg LC, Poole AR, et al. Cartilage and bone metabolism in rheumatoid arthritis. Differences between rapid and slow progression of disease identified by serum markers of cartilage metabolism. *J. Clin. Invest.* 1995 Mar;95(3):1071–7.
  78. Charni N, Juillet F, Garnero P. Urinary type II collagen helical peptide (HELIX-II) as a new biochemical marker of cartilage degradation in patients with osteoarthritis and rheumatoid arthritis. *Arthritis Rheum.* 2005;52(4):1081–90.
  79. Saxne T, Zunino L, Heinegård D. Increased release of bone sialoprotein into synovial fluid reflects tissue destruction in rheumatoid arthritis. *Arthritis Rheum.* 1995 Jan;38(1):82–90.
  80. Aman S, Risteli J, Luukkainen R, Risteli L, Kauppi M, Nieminen P, et al. The value of synovial fluid analysis in the assessment of knee joint destruction in arthritis in a three year follow up study. *Annals of the Rheumatic Diseases.* 1999 Sep;58(9):559–62.
  81. Hakala M, Risteli J, Aman S, Kautiainen H, Korpela M, Hannonen P, et al. Combination drug strategy in recent-onset rheumatoid arthritis suppresses collagen I degradation and is associated with retardation of radiological progression. *Scand J Rehabil Med.* 2008 Jan;37(2):90–3.
  82. Rioja I, Hughes FJ, Sharp CH, Warnock LC, Montgomery DS, Akil M, et al. Potential novel biomarkers of disease activity in rheumatoid arthritis patients: CXCL13, CCL23, transforming growth factor  $\alpha$ , tumor necrosis factor receptor superfamily member 9, and macrophage colony-stimulating factor. *Arthritis Rheum.* 2008 Aug;58(8):2257–67.
  83. Young-Min S, Cawston T, Marshall N, Coady D, Christgau S, Saxne T, et al. Biomarkers predict radiographic progression in early rheumatoid arthritis and perform well compared with traditional markers. *Arthritis Rheum.* 2007;56(10):3236–47.
  84. Rantapää-Dahlqvist S, de Jong BAW, Berglin E, Hallmans G, Wadell G, Stenlund H, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. *Arthritis Rheum.* 2003 Oct;48(10):2741–9.
  85. Majka DS, Deane KD, Parrish LA, Lazar AA, Barón AE, Walker CW, et al. Duration of preclinical rheumatoid arthritis-related autoantibody positivity increases in subjects with older age at time of disease diagnosis. *Annals of the Rheumatic Diseases.* 2008 Jun;67(6):801–7.
  86. Bos WH, Wolbink GJ, BOERS M, Tjihuis GJ, de Vries N, van der Horst-Bruinsma IE, et al. Arthritis development in patients with arthralgia is strongly associated with anti-citrullinated protein antibody status: a prospective cohort study. *Annals of the Rheumatic Diseases.* 2010 Mar 11;69(3):490–4.

87. van de Stadt LA, de Koning MHMT, van de Stadt RJ, Wolbink G, Dijkmans BAC, Hamann D, et al. Development of the anti-citrullinated protein antibody repertoire prior to the onset of rheumatoid arthritis. *Arthritis Rheum.* 2011 Oct 28;63(11):3226–33.
88. Sokolove J, Bromberg R, Deane KD, Lahey LJ, Derber LA, Chandra PE, et al. Autoantibody Epitope Spreading in the Pre-Clinical Phase Predicts Progression to Rheumatoid Arthritis. Matloubian M, editor. *PLoS ONE.* 2012 May 25;7(5):e35296.
89. Ioan-Facsinay A, Willemze A, Robinson DB, Peschken CA, Markland J, van der Woude D, et al. Marked differences in fine specificity and isotype usage of the anti-citrullinated protein antibody in health and disease. *Arthritis Rheum.* 2008 Oct;58(10):3000–8.
90. Lundberg K, Bengtsson C, Kharlamova N, Reed E, Jiang X, Kallberg H, et al. Genetic and environmental determinants for disease risk in subsets of rheumatoid arthritis defined by the anticitrullinated protein/peptide antibody fine specificity profile. *Annals of the Rheumatic Diseases.* 2013 May;72(5):652–8.
91. Silman AJ, Newman J, MacGregor AJ. Cigarette smoking increases the risk of rheumatoid arthritis. Results from a nationwide study of disease-discordant twins. *Arthritis Rheum.* 1996 May;39(5):732–5.
92. Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J, et al. A new model for an etiology of rheumatoid arthritis: Smoking may trigger HLA–DR (shared epitope)–restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum.* 2005;54(1):38–46.
93. Padyukov L, Silva C, Stolt P, Alfredsson L, Klareskog L, for the Epidemiological Investigation of Rheumatoid Arthritis Study Group. A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. *Arthritis Rheum.* 2004 Oct 8;50(10):3085–92.
94. Wegner N, Wait R, Sroka A, Eick S, Nguyen K-A, Lundberg K, et al. Peptidylarginine deiminase from *Porphyromonas gingivalis* citrullinates human fibrinogen and  $\alpha$ -enolase: Implications for autoimmunity in rheumatoid arthritis. *Arthritis Rheum.* 2010 May 6;62(9):2662–72.
95. HITCHON CA, CHANDAD F, FERUCCI ED, Willemze A, Ioan-Facsinay A, van der WOUDE D, et al. Antibodies to *Porphyromonas gingivalis* Are Associated with Anticitrullinated Protein Antibodies in Patients with Rheumatoid Arthritis and Their Relatives. *The Journal of Rheumatology.* 2010 Jun 1;37(6):1105–12.
96. de Pablo P, Dietrich T, Chapple I, Milward M, Chowdhury M et al. The auto-antibody repertoire in periodontitis: a role in the induction of autoimmunity to citrillunated proteins in rheumatoid arthritis? *Annals of the Rheumatic Diseases.* 2013; 73: 580-586
97. Makrygiannakis D, Hermansson M, Ulfgren A-K, Nicholas AP, Zendman AJW, Eklund A, et al. Smoking increases peptidylarginine deiminase 2 enzyme expression in human lungs and increases citrullination in BAL cells. *Annals of the Rheumatic Diseases.* 2008 Oct;67(10):1488–92.
98. Demoruelle MK, Weisman MH, Simonian PL, Lynch DA, Sachs PB, Pedraza IF, et al. Brief Report: Airways abnormalities and rheumatoid arthritis-related autoantibodies in subjects without arthritis: Early injury or initiating site of autoimmunity? *Arthritis Rheum.* 2012 May 25;64(6):1756–61.
99. van de Sande MGH, de Hair MJH, van der Leij C, Klarenbeek PL, Bos WH, Smith MD, et al. Different stages of rheumatoid arthritis: features of the synovium in the preclinical phase. *Annals of the Rheumatic Diseases.* 2011 Apr 3;70(5):772–7.
100. Karlson EW, Deane K. Environmental and Gene-Environment Interactions and Risk

of Rheumatoid Arthritis. *Rheumatic Disease Clinics of North America*. 2012 May;38(2):405–26.

101. Detert J, Pischon N, Burmester GR, Buttgereit F. The association between rheumatoid arthritis and periodontal disease. *Arthritis Res Ther*. 2010;12(5):218.
102. Linos A, Worthington JW, O'Fallon WM, Kurland LT. The epidemiology of rheumatoid arthritis in Rochester, Minnesota: a study of incidence, prevalence, and mortality. *Am. J. Epidemiol*. 1980 Jan;111(1):87–98.
103. Silman A, Kay A, Brennan P. Timing of pregnancy in relation to the onset of rheumatoid arthritis. *Arthritis Rheum*. 1992 Feb;35(2):152–5.
104. Goemaere S, Ackerman C, Goethals K, De Keyser F, Van der Straeten C, Verbruggen G, et al. Onset of symptoms of rheumatoid arthritis in relation to age, sex and menopausal transition. *The Journal of Rheumatology*. 1990 Dec;17(12):1620–2.
105. Pedersen M, Jacobsen S, Klarlund M, Pedersen BV, Wiik A, Wohlfahrt J, et al. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis Res Ther*. 2006;8(4):R133.
106. Bhatia SS, Majka DS, Kittelson JM, Parrish LA, FERUCCI ED, Deane KD, et al. Rheumatoid factor seropositivity is inversely associated with oral contraceptive use in women without rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 2006 Aug 11;66(2):267–9.
107. Voigt LF, Koepsell TD, Nelson JL, Dugowson CE, Daling JR. Smoking, obesity, alcohol consumption, and the risk of rheumatoid arthritis. *Epidemiology*. 1994 Sep;5(5):525–32.
108. Symmons DP, Bankhead CR, Harrison BJ, Brennan P, Barrett EM, Scott DG, et al. Blood transfusion, smoking, and obesity as risk factors for the development of rheumatoid arthritis: results from a primary care-based incident case-control study in Norfolk, England. *Arthritis Rheum*. 1997 Nov;40(11):1955–61.
109. Wesley A, Bengtsson C, Elkan A-C, Klareskog L, Alfredsson L, Wedrén S, et al. Association between body mass index and anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis: Results from a population-based case-control study. *Arthritis Care Res*. 2012 Dec 27;65(1):107–12.
110. Källberg H, Jacobsen S, Bengtsson C, Pedersen M, Padyukov L, Garred P, et al. Alcohol consumption is associated with decreased risk of rheumatoid arthritis: results from two Scandinavian case-control studies. *Annals of the Rheumatic Diseases*. 2009 Feb;68(2):222–7.
111. Smolen JS, Landewe R, Breedveld FC, Dougados M, Emery P, Gaujoux-Viala C, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. *Annals of the Rheumatic Diseases*. 2010 May 24;69(6):964–75.
112. Saag KG, Teng GG, Patkar NM, Anuntiyo J, Finney C, Curtis JR, et al. American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. *Arthritis Rheum*. 2008 Jun 15;59(6):762–84.
113. Quinn MA, Emery P. Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. *Clin. Exp. Rheumatol*. 2003 Aug;21(5 Suppl 31):S154–7.
114. Korpela M, Laasonen L, Hannonen P, Kautiainen H, Leirisalo-Repo M, Hakala M, et al. Retardation of joint damage in patients with early rheumatoid arthritis by initial

- aggressive treatment with disease-modifying antirheumatic drugs: Five-year experience from the FIN-RACo study. *Arthritis Rheum.* 2004;50(7):2072–81.
115. Rantalaiho V, Korpela M, Hannonen P, Kautiainen H, Järvenpää S, Leirisalo-Repo M, et al. The good initial response to therapy with a combination of traditional disease-modifying antirheumatic drugs is sustained over time: The eleven-year results of the Finnish rheumatoid arthritis combination therapy trial. *Arthritis Rheum.* 2009 May;60(5):1222–31.
  116. Nell VPK. Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis. *Rheumatology.* 2004 May 4;43(7):906–14.
  117. van der Linden MPM, le Cessie S, Raza K, van der Woude D, Knevel R, Huizinga TWJ, et al. Long-term impact of delay in assessment of patients with early arthritis. *Arthritis Rheum.* 2010 Nov 30;62(12):3537–46.
  118. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *The Lancet.* 2004 Jul;364(9430):263–9.
  119. Verstappen SMM, Jacobs JWG, van der Veen MJ, Heurkens AHM, Schenk Y, Borg ter EJ, et al. Intensive treatment with methotrexate in early rheumatoid arthritis: aiming for remission. Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA, an open-label strategy trial). *Annals of the Rheumatic Diseases.* 2007 Nov 1;66(11):1443–9.
  120. Mottonen T, Hannonen P, Leirisalo-Repo M, Nissilä M, Kautiainen H, Korpela M, et al. Comparison of combination therapy with single-drug therapy in early rheumatoid arthritis: a randomised trial. FIN-RACo trial group. *Lancet.* 1999 May 8;353(9164):1568–73.
  121. Brown AK, Quinn MA, Karim Z, Conaghan PG, Peterfy CG, Hensor E, et al. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: Evidence from an imaging study may explain structural progression. *Arthritis Rheum.* 2006;54(12):3761–73.
  122. Wakefield RJ. Should oligoarthritis be reclassified? Ultrasound reveals a high prevalence of subclinical disease. *Annals of the Rheumatic Diseases.* 2004 Apr 1;63(4):382–5.
  123. Goekoop-Ruiterman YPM, de Vries-Bouwstra JK, Allaart CF, van Zeben D, Kerstens PJSM, Hazes JMW, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): A randomized, controlled trial. *Arthritis Rheum.* 2005;52(11):3381–90.
  124. Smolen JS, ALETAHA D, Bijlsma JWJ, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. *Annals of the Rheumatic Diseases.* 2010 Mar 17;69(4):631–7.
  125. RHEUMATOID ARTHRITIS. National clinical guideline for management and treatment in adults. 2012 Feb 22;:1–234.
  126. Cohen G, Gossec L, Dougados M, Cantagrel A, Goupille P, Daures J, et al. Radiological damage in patients with rheumatoid arthritis on sustained remission. *Annals of the Rheumatic Diseases.* 2007 Mar 1;66(3):358–63.
  127. Smolen JS, Han C, van der Heijde DMFM, Emery P, BATHON JM, Keystone E, et al. Radiographic changes in rheumatoid arthritis patients attaining different disease activity states with methotrexate monotherapy and infliximab plus methotrexate: the impacts of remission and tumour necrosis factor blockade. *Annals of the Rheumatic Diseases.* 2009 May 12;68(6):823–7.

128. Fransen J. Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial. *Annals of the Rheumatic Diseases*. 2005 Sep 1;64(9):1294–8.
129. Wakefield RJ, d'Agostino MA, Naredo E, Buch MH, Iagnocco A, Terslev L, et al. After treat-to-target: can a targeted ultrasound initiative improve RA outcomes? *Annals of the Rheumatic Diseases*. 2012 May 7;71(6):799–803.
130. Khanna D, Oh M, Furst DE, Ranganath V, Gold RH, Sharp JT, et al. Evaluation of the preliminary definitions of minimal disease activity and remission in an early seropositive rheumatoid arthritis cohort. *Arthritis Rheum*. 2007;57(3):440–7.
131. O'Dell JR, Leff R, Paulsen G, Haire C, Mallek J, Eckhoff PJ, et al. Treatment of rheumatoid arthritis with methotrexate and hydroxychloroquine, methotrexate and sulfasalazine, or a combination of the three medications: Results of a two-year, randomized, double-blind, placebo-controlled trial. *Arthritis Rheum*. 2002 May 8;46(5):1164–70.
132. Tugwell P, Pincus T, Yocum D, Stein M, Gluck O, Kraag G, et al. Combination therapy with cyclosporine and methotrexate in severe rheumatoid arthritis. The Methotrexate-Cyclosporine Combination Study Group. *N Engl J Med*. 1995 Jul 20;333(3):137–41.
133. Stein CM, Pincus T, Yocum D, Tugwell P, Wells G, Gluck O, et al. Combination treatment of severe rheumatoid arthritis with cyclosporine and methotrexate for forty-eight weeks: an open-label extension study. The Methotrexate-Cyclosporine Combination Study Group. *Arthritis Rheum*. 1997 Oct;40(10):1843–51.
134. Kremer JM, Genovese MC, Cannon GW, Caldwell JR, Cush JJ, Furst DE, et al. Concomitant leflunomide therapy in patients with active rheumatoid arthritis despite stable doses of methotrexate. A randomized, double-blind, placebo-controlled trial. *Ann. Intern. Med*. 2002 Nov 5;137(9):726–33.
135. Lehman AJ, Esdaile JM, Klinkhoff AV, Grant E, Fitzgerald A, Canvin J, et al. A 48-week, randomized, double-blind, double-observer, placebo-controlled multicenter trial of combination methotrexate and intramuscular gold therapy in rheumatoid arthritis: Results of the METGO study. *Arthritis Rheum*. 2005 May;52(5):1360–70.
136. Rachapalli SM, Williams R, Walsh DA, Young A, Kiely PDW, Choy EH, et al. First-line DMARD choice in early rheumatoid arthritis--do prognostic factors play a role? *Rheumatology*. 2010 Jun 16;49(7):1267–71.
137. Donahue K, Gartlehner G, Jonas D, Lux L, Thieda P, Jonas B, et al. Systematic Review: Comparative Effectiveness and Harms of Disease-Modifying Medications for Rheumatoid Arthritis. *Ann. Intern. Med*. 2008 Jan 15;148(2):124–34.
138. Katchamart W, Trudeau J, Phumethum V, Bombardier C. Methotrexate monotherapy versus methotrexate combination therapy with non-biologic disease modifying anti-rheumatic drugs for rheumatoid arthritis (Review). *Cochrane Database of Systematic Reviews*. 2010 Jan 1;(4):1–111.
139. Rezaei H, Saevarsdottir S, Forslind K, Albertsson K, Wallin H, Bratt J, et al. In early rheumatoid arthritis, patients with a good initial response to methotrexate have excellent 2-year clinical outcomes, but radiological progression is not fully prevented: data from the methotrexate responders population in the SWEFOT trial. *Annals of the Rheumatic Diseases*. 2011 Dec 28;71(2):186–91.
140. Hetland ML, Stengaard-Pedersen K, Junker P, Lottenburger T, Ellingsen T, Andersen LS, et al. Combination treatment with methotrexate, cyclosporine, and intraarticular betamethasone compared with methotrexate and intraarticular betamethasone in early active rheumatoid arthritis: An investigator-initiated, multicenter, randomized, double-blind, parallel-group, placebo-controlled study. *Arthritis Rheum*. 2006;54(5):1401–9.

141. Haagsma CJ, van Riel PL, de Jong AJ, van De Putte LB. Combination of sulphasalazine and methotrexate versus the single components in early rheumatoid arthritis: a randomized, controlled, double-blind, 52 week clinical trial. *Br. J. Rheumatol.* 1997 Oct;36(10):1082–8.
142. Dougados M, Combe B, Cantagrel A, Goupille P, Olive P, Schattenkirchner M, et al. Combination therapy in early rheumatoid arthritis: a randomised, controlled, double blind 52 week clinical trial of sulphasalazine and methotrexate compared with the single components. *Annals of the Rheumatic Diseases.* 1999 Apr 1;58(4):220–5.
143. BOERS M, Verhoeven AC, Markusse HM, van de Laar MA, Westhovens R, van Denderen JC, et al. Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. *Lancet.* 1997 Aug 2;350(9074):309–18.
144. Choy EHS. A meta-analysis of the efficacy and toxicity of combining disease-modifying anti-rheumatic drugs in rheumatoid arthritis based on patient withdrawal. *Rheumatology.* 2005 Jun 21;44(11):1414–21.
145. Goekoop-Ruiterman YPM, De Vries-Bouwstra JK, Allaart CF, van Zeben D, Kerstens PJSM, Hazes JMW, et al. Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial. *Ann. Intern. Med.* 2007 Mar 20;146(6):406–15.
146. Saunders SA, Capell HA, Stirling A, Vallance R, Kincaid W, McMahon AD, et al. Triple therapy in early active rheumatoid arthritis: A randomized, single-blind, controlled trial comparing step-up and parallel treatment strategies. *Arthritis Rheum.* 2008;58(5):1310–7.
147. van Vollenhoven R, Ernestam S, Geborek P, Petersson I, Cöster L, Waltbrand E, et al. Addition of infliximab compared with addition of sulfasalazine and hydroxychloroquine to methotrexate in patients with early rheumatoid arthritis (Swefot trial): 1-year results of a randomised trial. *The Lancet.* Elsevier Ltd; 2009 Aug 8;374(9688):459–66.
148. O'dell J, Curtis J, Cofield S, Mikuls T, Moreland L. Validation Of Methotrexate First Strategy In Early Rheumatoid Arthritis: A Randomized, Double-Blind, 2-Year Trial. *Arthritis Rheum.* 2011. p. 1696.
149. Moreland LW, O'Dell JR, Paulus HE, Curtis JR, Bathon JM, St Clair EW, et al. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: The Treatment of Early Aggressive Rheumatoid Arthritis trial. *Arthritis Rheum* [Internet]. <http://dx.doi.org/10.1002/art.34498>
150. Szkudlarek M, Klarlund M, Narvestad E, Court-Payen M, Strandberg C, Jensen KE, et al. Ultrasonography of the metacarpophalangeal and proximal interphalangeal joints in rheumatoid arthritis: a comparison with magnetic resonance imaging, conventional radiography and clinical examination. *Arthritis Res Ther.* 2006;8(2):R52.
151. Filer A, de Pablo P, Allen G, Nightingale P, Jordan A, Jobanputra P, et al. Utility of ultrasound joint counts in the prediction of rheumatoid arthritis in patients with very early synovitis. *Annals of the Rheumatic Diseases.* 2011 Feb 3;70(3):500–7.
152. Salaffi F, Filippucci E, Carotti M, Naredo E, Meenagh G, Ciapetti A, et al. Inter-observer agreement of standard joint counts in early rheumatoid arthritis: a comparison with grey scale ultrasonography a preliminary study. *Rheumatology.* 2008 Jan 1;47(1):54–8.
153. STONE MA, WHITE LM, GLADMAN DD, INMAN RD, CHAYA S, LAX M, et al. Significance of Clinical Evaluation of the Metacarpophalangeal Joint in Relation to Synovial/Bone Pathology in Rheumatoid and Psoriatic Arthritis Detected by

Magnetic Resonance Imaging. *The Journal of Rheumatology*. 2009 Dec 4;36(12):2751–7.

154. Wolfe F, Michaud K. Severe rheumatoid arthritis (RA), worse outcomes, comorbid illness, and sociodemographic disadvantage characterize ra patients with fibromyalgia. *The Journal of Rheumatology*. 2004 Apr 1;31(4):695–700.
155. Pollard LC, Kingsley GH, Choy EH, Scott DL. Fibromyalgic rheumatoid arthritis and disease assessment. *Rheumatology*. 2010 Apr 13;49(5):924–8.
156. Wolfe F, Michaud K. The clinical and research significance of the erythrocyte sedimentation rate. *Journal of Rheumatology*. 1994 Jul 1;21(7):1227–37.
157. Pincus T. Advantages and limitations of quantitative measures to assess rheumatoid arthritis: joint counts, radiographs, laboratory tests, and patient questionnaires. *Bull NYU Hosp Jt Dis*. 2006;64(1-2):32–9.
158. Dixon JS, Bird HA, Sitton NG, Pickup ME, Wright V. C-reactive protein in the serial assessment of disease activity in rheumatoid arthritis. *Scand J Rehabil Med*. 1984;13(1):39–44.
159. van der Heijde DM, van't Hof MA, van Riel PL, van Leeuwen MA, van Rijswijk MH, van De Putte LB. Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 1992 Feb;51(2):177–81.
160. Moreland LW, Baumgartner SW, Schiff MH, Tindall EA, Fleischmann RM, Weaver AL, et al. Treatment of rheumatoid arthritis with a recombinant human tumor necrosis factor receptor (p75)-Fc fusion protein. *N Engl J Med*. 1997 Jul 17;337(3):141–7.
161. Anderson JK, Zimmerman L, Caplan L, Michaud K. Measures of rheumatoid arthritis disease activity: Patient (PtGA) and Provider (PrGA) Global Assessment of Disease Activity, Disease Activity Score (DAS) and Disease Activity Score With 28-Joint Counts (DAS28), Simplified Disease Activity Index (SDAI), CI. *Arthritis Care Res*. 2011 Nov 7;63(S11):S14–S36.
162. Madsen R, Lundstedt T, Trygg J. Chemometrics in metabolomics--a review in human disease diagnosis. *Anal. Chim. Acta*. 2010 Feb 5;659(1-2):23–33.
163. Fransen J, Stucki G, van Riel PLCM. Rheumatoid arthritis measures: Disease Activity Score (DAS), Disease Activity Score-28 (DAS28), Rapid Assessment of Disease Activity in Rheumatology (RADAR), and Rheumatoid Arthritis Disease Activity Index (RADAI). *Arthritis Rheum*. 2003 Oct 15;49(S5):S214–24.
164. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI, et al. A Trial of Etanercept, a Recombinant Tumor Necrosis Factor Receptor:Fc Fusion Protein, in Patients with Rheumatoid Arthritis Receiving Methotrexate. *N Engl J Med*. 1999 Jan 28;340(4):253–9.
165. Anderson J, Caplan L, Yazdany J, Robbins ML, Neogi T, Michaud K, et al. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. *Arthritis Care Res*. 2012 Apr 2;64(5):640–7.
166. Klareskog L, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, et al. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. *The Lancet*. Elsevier; 2004;363(9410):675–81.
167. van der Heijde D, Burmester G, Melo-Gomes J, Codreanu C, Mola EM, Pedersen R, et al. The safety and efficacy of adding etanercept to methotrexate or methotrexate to etanercept in moderately active rheumatoid arthritis patients

previously treated with monotherapy. 2008.

168. Luqmani R. British Society for Rheumatology and British Health Professionals in Rheumatology Guideline for the Management of Rheumatoid Arthritis (The first 2 years). *Rheumatology*. 2006 Apr 25.
169. Combe B, Landewe R, Lukas C, Bolosiu HD, Breedveld F, Dougados M, et al. EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Annals of the Rheumatic Diseases*. 2006 Jun 30;66(1):34–45.
170. NICE. NICE clinical guideline 79. 2010 Sep 9;:1–35.
171. Wolfe F, Michaud K, Pincus T. A composite disease activity scale for clinical practice, observational studies, and clinical trials: the patient activity scale (PAS/PAS-II). *Journal of Rheumatology*. 2005 Dec 1;32(12):2410–5.
172. Shaver T, Anderson J, Weidensaul D, Shadi S, Busch R, Mikuls T, et al. The Problem of Rheumatoid Arthritis Disease Activity and Remission in Clinical Practice. *Journal of Rheumatology*. 2008 Jun 1;35(6):1015–22.
173. Pincus T, Swearingen CJ, Bergman M, Yazici Y. RAPID3 (Routine Assessment of Patient Index Data 3), a Rheumatoid Arthritis Index Without Formal Joint Counts for Routine Care: Proposed Severity Categories Compared to Disease Activity Score and Clinical Disease Activity Index Categories. *The Journal of Rheumatology*. 2008 Sep 15;35(11):2136.
174. Pincus T, Swearingen CJ, Bergman MJ, Colglazier CL, Kaell AT, Kunath AM, et al. RAPID3 (Routine Assessment of Patient Index Data) on an MDHAQ (Multidimensional Health Assessment Questionnaire): Agreement with DAS28 (Disease Activity Score) and CDAI (Clinical Disease Activity Index) activity categories, scored in five versus more than ninety seconds. *Arthritis Care Res*. 2010 Feb;62(2):181–9.
175. RINTELEN B, HAINDL PM, SAUTNER J, LEEB BA, DEUTSCH C, Leeb BF. The Rheumatoid Arthritis Disease Activity Index-5 in Daily Use. Proposal for Disease Activity Categories. *The Journal of Rheumatology*. 2009 May 12;36(5):918–24.
176. Leeb BF, Haindl PM, Maktari A, Nothnagl T, Rintelen B. Patient-centered rheumatoid arthritis disease activity assessment by a modified RADAI. *The Journal of Rheumatology*. 2008 Jul;35(7):1294–9.
177. Fransen J, Häuselmann H, Michel BA, Caravatti M, Stucki G. Responsiveness of the self-assessed rheumatoid arthritis disease activity index to a flare of disease activity. *Arthritis Rheum*. 2001 Jan;44(1):53–60.
178. Aletaha D, Nell VPK, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther*. 2005;7(4):R796–806.
179. ALETAHA D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin. Exp. Rheumatol*. 2005 Aug;23(5 Suppl 39):S100–8.
180. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van De Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum*. 1995 Jan;38(1):44–8.
181. Hensor EMA, Emery P, Bingham SJ, Conaghan PG, the YEAR Consortium. Discrepancies in categorizing rheumatoid arthritis patients by DAS-28(ESR) and

- DAS-28(CRP): can they be reduced? *Rheumatology*. 2010 Jul 13;49(8):1521–9.
182. Smolen JS, Aletaha D. Interleukin-6 receptor inhibition with tocilizumab and attainment of disease remission in rheumatoid arthritis: The role of acute-phase reactants. *Arthritis Rheum*. 2010 Dec 28;63(1):43–52.
183. Drossaers-Bakker KW, de Buck M, van Zeben D, Zwinderman AH, Breedveld FC, Hazes JM. Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time. *Arthritis Rheum*. 1999 Sep;42(9):1854–60.
184. Smolen JS. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology*. 2003 Feb 1;42(2):244–57.
185. RINTELEN B, Haindl P, Maktari A, Nothnagl T, Hartl E, Leeb B. SDAI/CDAI levels in rheumatoid arthritis patients are highly dependent on patient's pain perception and gender. *Scand J Rehabil Med*. 2008 Nov 1;37(6):410–3.
186. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LHD, Funovits J, et al. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis Rheum*. 2011 Feb 3;63(3):573–86.
187. Soubrier M, Zerkak D, Gossec L, Ayrat X, Roux C, Dougados M. Which variables best predict change in rheumatoid arthritis therapy in daily clinical practice? *The Journal of Rheumatology*. 2006 Jul;33(7):1243–6.
188. Aletaha D, Ward MM, Machold KP, Nell VPK, Stamm T, Smolen JS. Remission and active disease in rheumatoid arthritis: Defining criteria for disease activity states. *Arthritis Rheum*. 2005;52(9):2625–36.
189. Mäkinen H, Kautiainen H, Hannonen P, Sokka T. Is DAS28 an appropriate tool to assess remission in rheumatoid arthritis? *Annals of the Rheumatic Diseases*. 2005 Oct;64(10):1410–3.
190. van der Heijde D, Klareskog L, BOERS M, Landewe R, Codreanu C, Bolosiu HD, et al. Comparison of different definitions to classify remission and sustained remission: 1 year TEMPO results. *Annals of the Rheumatic Diseases*. 2005 Nov;64(11):1582–7.
191. ALETAHA D, Smolen J. Remission of rheumatoid arthritis: should we care about definitions? *Clin. Exp. Rheumatol. CLINICAL AND EXPERIMENTAL RHEUMATOLOGY SAS*; 2006;24(6):45.
192. Aletaha D, Smolen JS. Joint damage in rheumatoid arthritis progresses in remission according to the Disease Activity Score in 28 joints and is driven by residual swollen joints. *Arthritis Rheum*. 2011 Nov 29;63(12):3702–11.
193. Molenaar ETH, Voskuyl AE, Dinant HJ, Bezemer PD, Boers M, Dijkmans BAC. Progression of radiologic damage in patients with rheumatoid arthritis in clinical remission. *Arthritis Rheum*. 2004;50(1):36–42.
194. van Tuyl LHD, Lems WF, Voskuyl AE, Kerstens PJSM, GARNERO P, Dijkmans BAC, et al. Tight control and intensified COBRA combination treatment in early rheumatoid arthritis: 90% remission in a pilot trial. *Annals of the Rheumatic Diseases*. 2007 Nov 29;67(11):1574–7.
195. Stenger A. Early effective suppression of inflammation in rheumatoid arthritis reduces radiographic progression. *Rheumatology*. 1998 Nov 1;37(11):1157–63.
196. Edmonds J, Lassere M, Sharp J, Bird P. Objectives Study in RA (OSRA): A RCT defining the best clinical target for disease activity control in RA. *Annals of the Rheumatic Diseases*. 2007. p. 325.

197. Kane D, Balint PV, Sturrock RD. Ultrasonography is superior to clinical examination in the detection and localization of knee joint effusion in rheumatoid arthritis. *The Journal of Rheumatology*. 2003 May;30(5):966–71.
198. Wakefield RJ, Freeston JE, Hensor EMA, Bryer D, Quinn MA, Emery P. Delay in imaging versus clinical response: A rationale for prolonged treatment with anti-tumor necrosis factor medication in early rheumatoid arthritis. *Arthritis Rheum*. 2007;57(8):1564–7.
199. Balsa A, de Miguel E, Castillo C, Peiteado D, Martin-Mola E. Superiority of SDAI over DAS-28 in assessment of remission in rheumatoid arthritis patients using power Doppler ultrasonography as a gold standard. *Rheumatology*. 2010 Mar 15;49(4):683–90.
200. Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG, et al. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum*. 2008 Oct;58(10):2958–67.
201. Brigden ML. Clinical utility of the erythrocyte sedimentation rate. *Am Fam Physician*. 1999 Oct 1;60(5):1443–50.
202. Pepys MB, Hirschfield GM. C-reactive protein: a critical update. *J. Clin. Invest*. 2003 Jun 15;111(12):1805–12.
203. Welsing PMJ, Fransen J, van Riel PLCM. Is the disease course of rheumatoid arthritis becoming milder?: Time trends since 1985 in an inception cohort of early rheumatoid arthritis. *Arthritis Rheum*. 2005;52(9):2616–24.
204. Leeb BF. The DAS28 in rheumatoid arthritis and fibromyalgia patients. *Rheumatology*. 2004 Jul 13;43(12):1504–7.
205. Taylor P, Manger B, Alvaro-Gracia J, Johnstone R, Gomez-Reino J, Eberhardt E, et al. Patient perceptions concerning pain management in the treatment of rheumatoid arthritis. *The Journal of international medical research*. Field House Publishing; 2010;38(4):1213–24.
206. Ranzolin A, Brenol JCT, Bredemeier M, Guarienti J, Rizzatti M, Feldman D, et al. Association of concomitant fibromyalgia with worse disease activity score in 28 joints, health assessment questionnaire, and short form 36 scores in patients with rheumatoid arthritis. *Arthritis Rheum*. 2009 Jun 15;61(6):794–800.
207. Lee YC, Cui J, Lu B, Frits ML, Iannaccone CK, Shadick NA, et al. Pain persists in DAS28 rheumatoid arthritis remission but not in ACR/EULAR remission: a longitudinal observational study. *Arthritis Res Ther*. 2011;13(3):R83.
208. McWilliams DF, Zhang W, Mansell JS, Kiely PD, Young A, Walsh DA. Predictors of change in bodily pain in early rheumatoid arthritis: The eran study. *Arthritis Care Res*. 2012;:n/a–n/a.
209. Leffler A-S, Kosek E, Lerndal T, Nordmark B, Hansson P. Somatosensory perception and function of diffuse noxious inhibitory controls (DNIC) in patients suffering from rheumatoid arthritis. *European Journal of Pain*. 2012 Jan 9;6(2):161–76.
210. FITZGERALD O, Soden M, Yanni G, Robinson R, Bresnihan B. Morphometric analysis of blood vessels in synovial membranes obtained from clinically affected and unaffected knee joints of patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 1991 Nov;50(11):792–6.
211. Schmidt WA, Völker L, Zacher J, Schläfke M, Ruhnke M, Gromnica-Ihle E. Colour Doppler ultrasonography to detect pannus in knee joint synovitis. *Clin. Exp. Rheumatol*. 2000 Jun;18(4):439–44.

212. Walther M, Harms H, Krenn V, Radke S, Faehndrich TP, Gohlke F. Correlation of power Doppler sonography with vascularity of the synovial tissue of the knee joint in patients with osteoarthritis and rheumatoid arthritis. *Arthritis Rheum.* 2001 Feb;44(2):331–8.
213. Koski JM, Saarakkala S, Helle M, Hakulinen U, Heikkinen JO, Hermunen H. Power Doppler ultrasonography and synovitis: correlating ultrasound imaging with histopathological findings and evaluating the performance of ultrasound equipments. *Annals of the Rheumatic Diseases.* 2006 May 25;65(12):1590–5.
214. Terslev L, Torp-Pedersen S, Savnik A, Recke von der P, Qvistgaard E, Danneskiold-Samsøe B, et al. Doppler ultrasound and magnetic resonance imaging of synovial inflammation of the hand in rheumatoid arthritis: A comparative study. *Arthritis Rheum.* 2003 Sep 11;48(9):2434–41.
215. Klauser A, Frauscher F, Schirmer M, Halpern E, Pallwein L, Herold M, et al. The value of contrast-enhanced color doppler ultrasound in the detection of vascularization of finger joints in patients with rheumatoid arthritis. *Arthritis Rheum.* 2002 Mar 7;46(3):647–53.
216. Peluso G, Michelutti A, Bosello S, Gremese E, Toluoso B, Ferraccioli G. Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 2010 Dec 15;70(1):172–5.
217. Scire CA, Montecucco C, Codullo V, Epis O, Todoerti M, Caporali R. Ultrasonographic evaluation of joint involvement in early rheumatoid arthritis in clinical remission: power Doppler signal predicts short-term relapse. *Rheumatology.* 2009 Aug 17;48(9):1092–7.
218. Saleem B, Brown A, Quinn M. Prediction of flare and long term outcome in DMARD-treated RA patients in remission: the value of imaging and new remission criteria. *Ann Rheum Dis.* 2011;70(Suppl 3):88.
219. Hau M, Kneitz C, Tony H-P, Keberle M, Jahns R, Jenett M. High resolution ultrasound detects a decrease in pannus vascularisation of small finger joints in patients with rheumatoid arthritis receiving treatment with soluble tumour necrosis factor alpha receptor (etanercept). *Annals of the Rheumatic Diseases.* 2002 Jan;61(1):55–8.
220. Terslev L, Torp-Pedersen S, Qvistgaard E, Kristoffersen H, Røggind H, Danneskiold-Samsøe B, et al. Effects of treatment with etanercept (Enbrel, TNRF:Fc) on rheumatoid arthritis evaluated by Doppler ultrasonography. *Annals of the Rheumatic Diseases.* 2003 Feb;62(2):178–81.
221. Terslev L, Torp-Pedersen S, Qvistgaard E, Danneskiold-Samsøe B, Bliddal H. Estimation of inflammation by Doppler ultrasound: quantitative changes after intra-articular treatment in rheumatoid arthritis. *Annals of the Rheumatic Diseases.* BMJ Publishing Group Ltd and European League Against Rheumatism; 2003;62(11):1049–53.
222. Teh J. Power Doppler ultrasound of rheumatoid synovitis: quantification of therapeutic response. *British Journal of Radiology.* 2003 Dec 1;76(912):875–9.
223. Taylor PC, Steuer A, Gruber J, Cosgrove DO, Blomley MJK, Marsters PA, et al. Comparison of ultrasonographic assessment of synovitis and joint vascularity with radiographic evaluation in a randomized, placebo-controlled study of infliximab therapy in early rheumatoid arthritis. *Arthritis Rheum.* 2004;50(4):1107–16.
224. Ellegaard K, Torp-Pedersen S, Holm CC, Danneskiold-Samsøe B, Bliddal H. Ultrasound in finger joints: Findings in normal subjects and pitfalls in the diagnosis of synovial disease. *ULTRASCHALL IN DER MEDIZIN. RUDIGERSTR 14, D-70469 STUTTGART, GERMANY: GEORG THIEME VERLAG KG; 2007*

Aug;28(4):401–408).

225. Naredo E. Interobserver reliability in musculoskeletal ultrasonography: results from a “Teach the Teachers” rheumatologist course. *Annals of the Rheumatic Diseases*. 2006 Jan 1;65(1):14–9.
226. Naredo E, Rodríguez M, Campos C, Rodríguez-heredia JM, Medina JA, Giner E, et al. Validity, reproducibility, and responsiveness of a twelve-joint simplified power Doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. *Arthritis Care and Research*. 2008 {APR 15};59(4):515–522.
227. McInnes IB, O'Dell JR. State-of-the-art: rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 2010 Oct 18;69(11):1898–906.
228. Maas K, Chan S, Parker J, Slater A, Moore J, Olsen N, et al. Cutting edge: molecular portrait of human autoimmune disease. *J. Immunol*. 2002 Jul 1;169(1):5–9.
229. van der Pouw Kraan TCTM, van Gaalen FA, Huizinga TWJ, Pieterman E, Breedveld FC, Verweij CL. Discovery of distinctive gene expression profiles in rheumatoid synovium using cDNA microarray technology: evidence for the existence of multiple pathways of tissue destruction and repair. *Genes Immun*. 2003 Apr;4(3):187–96.
230. Olsen N. A gene expression signature for recent onset rheumatoid arthritis in peripheral blood mononuclear cells. *Annals of the Rheumatic Diseases*. 2004 Nov 1;63(11):1387–92.
231. Devauchelle V, Marion S, Cagnard N, Mistou S, Falgarone G, Breban M, et al. DNA microarray allows molecular profiling of rheumatoid arthritis and identification of pathophysiological targets. *Genes Immun*. 2004 Dec;5(8):597–608.
232. Batliwalla FM, Baechler EC, Xiao X, Li W, Balasubramanian S, Khalili H, et al. Peripheral blood gene expression profiling in rheumatoid arthritis. *Genes Immun*. 2005 Jun 23;6(5):388–97.
233. Szodoray P, Alex P, Frank MB, Turner M, Turner S, Knowlton N, et al. A genome-scale assessment of peripheral blood B-cell molecular homeostasis in patients with rheumatoid arthritis. *Rheumatology*. 2006 Aug 18;45(12):1466–76.
234. Lequerré T, Gauthier-Jauneau A-C, Bansard C, Derambure C, Hiron M, Vittecoq O, et al. Gene profiling in white blood cells predicts infliximab responsiveness in rheumatoid arthritis. *Arthritis Res Ther*. 2006;8(4):R105.
235. Lindberg J, af Klint E, Catrina AI, Nilsson P, Klareskog L, Ulfgren A-K, et al. Effect of infliximab on mRNA expression profiles in synovial tissue of rheumatoid arthritis patients. *Arthritis Res Ther*. 2006;8(6):R179.
236. van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Rustenburg F, Baggen JM, Verweij CL, et al. Responsiveness to anti-tumour necrosis factor  $\alpha$  therapy is related to pre-treatment tissue inflammation levels in rheumatoid arthritis patients. *Annals of the Rheumatic Diseases*. 2008.
237. Sekiguchi N, Kawauchi S, Furuya T, Inaba N, Matsuda K, Ando S, et al. Messenger ribonucleic acid expression profile in peripheral blood cells from RA patients following treatment with an anti-TNF- $\alpha$  monoclonal antibody, infliximab. *Rheumatology*. 2008 Mar 11;47(6):780–8.
238. Koczan D, Drynda S, Hecker M, Drynda A, Guthke R, Kekow J, et al. Molecular discrimination of responders and nonresponders to anti-TNF alpha therapy in rheumatoid arthritis by etanercept. *Arthritis Res Ther*. 2008;10(3):R50.
239. Stuhlmüller B, Häupl T, Hernandez MM, Grützkau A, Kuban R-J, Tandon N, et al.

CD11c as a Transcriptional Biomarker to Predict Response to Anti-TNF Monotherapy With Adalimumab in Patients With Rheumatoid Arthritis. *Clin Pharmacol Ther.* 2009 Dec 23;87(3):311–21.

240. van der Pouw Kraan TCTM, Wijbrandts CA, van Baarsen LGM, Voskuyl AE, Rustenburg F, Baggen JM, et al. Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. *Annals of the Rheumatic Diseases.* 2007 Jan 12;66(8):1008–14.
241. Lindberg J, Wijbrandts CA, van Baarsen LG, Nader G, Klareskog L, Catrina A, et al. The gene expression profile in the synovium as a predictor of the clinical response to infliximab treatment in rheumatoid arthritis. *PLoS ONE.* 2010;5(6):e11310.
242. Madsen RK, Lundstedt T, Gabrielsson J, Sennbro C-J, Alenius G-M, Moritz T, et al. Diagnostic properties of metabolic perturbations in rheumatoid arthritis. *Arthritis Res Ther.* 2011;13(1):R19.
243. Hügler T, Kovacs H, Heijnen I, Daikeler T. Synovial fluid metabolomics in different forms of arthritis assessed by nuclear magnetic resonance spectroscopy. *Clinical and Experimental Rheumatology.* 2012; 30 (2): 240-245
244. van Wietmarschen HA, Dai W, van der Kooij AJ, Reijmers TH, Schroën Y, Wang M, et al. Characterization of Rheumatoid Arthritis Subtypes Using Symptom Profiles, Clinical Chemistry and Metabolomics Measurements. Sawalha AH, editor. *PLoS ONE.* 2012 Sep 12;7(9):e44331.
245. Raza K, Breese M, Nightingale P, Kumar K, Potter T, Carruthers DM, et al. Predictive value of antibodies to cyclic citrullinated peptide in patients with very early inflammatory arthritis. *The Journal of Rheumatology.* 2005 Feb;32(2):231–8.
246. Karim Raza MRCP PD, Andrew Filer MRCP PD. Predicting the development of RA in patients with early undifferentiated arthritis. *Best Pract Res Clin Rheumatol.* Elsevier Ltd; 2009 Feb 1;23(1):25–36.
247. Meyer O, Labarre C, Dougados M, Goupille P, Cantagrel A, Dubois A, et al. Anticitrullinated protein/peptide antibody assays in early rheumatoid arthritis for predicting five year radiographic damage. *Annals of the Rheumatic Diseases.* 2003 Feb 1;62(2):120–6.
248. SIGN. SIGN Publication No. 48 - Management of early rheumatoid arthritis. 2001 Jan 26:1–57.
249. Stack RJ, Shaw K, Mallen C, Herron-Marx S, Horne R, Raza K. Delays in help seeking at the onset of the symptoms of rheumatoid arthritis: a systematic synthesis of qualitative literature. *Annals of the Rheumatic Diseases.* 2012 Apr;71(4):493–7.
250. Haavardsholm E, OSTERGAARD M. Reliability and sensitivity to change of the OMERACT rheumatoid arthritis magnetic resonance imaging score in a multireader, longitudinal setting. *Arthritis and Rheumatism.* 2005; 52(10): 3860-3867
251. Klarenbeek NB, Guler-Yuksel M, van der Kooij SM, Han KH, Roday HK, Kerstens PJSM, et al. The impact of four dynamic, goal-steered treatment strategies on the 5-year outcomes of rheumatoid arthritis patients in the BeSt study. *Annals of the Rheumatic Diseases.* 2011 May 2;70(6):1039–46.
252. van Vollenhoven RF, Geborek P, Forslind K, Albertsson K, Ernestam S, Petersson IF, et al. Conventional combination treatment versus biological treatment in methotrexate-refractory early rheumatoid arthritis: 2 year follow-up of the randomised, non-blinded, parallel-group Swefot trial. *Lancet.* 2012 Mar 28.
253. Van Everdingen AA, Jacobs JWJ, Siewertsz Van Reesema DR, Bijlsma JWJ. Low-dose prednisone therapy for patients with early active rheumatoid arthritis: clinical

- efficacy, disease-modifying properties, and side effects: a randomized, double-blind, placebo-controlled clinical trial. *Ann. Intern. Med.* 2002 Jan 1;136(1):1–12.
254. Svensson B, Boonen A, Albertsson K. Low-dose prednisolone in addition to the initial disease-modifying antirheumatic drug in patients with early active rheumatoid arthritis reduces joint destruction and increases the remission rate: A two-year randomized trial. *Arthritis & Rheumatism.* 2005; 52(11): 3360-3370
  255. Montecucco C, Todoerti M, Sakellariou G, Scirè CA, Caporali R. Low-dose oral prednisone improves clinical and ultrasonographic remission rates in early rheumatoid arthritis: results of a 12-month open-label randomised study. *Arthritis Res Ther.* 2012;14(3):R112.
  256. Kirwan JR. The effect of glucocorticoids on joint destruction in rheumatoid arthritis. The Arthritis and Rheumatism Council Low-Dose Glucocorticoid Study Group. *N Engl J Med.* 1995 Jul 20;333(3):142–6.
  257. Wassenberg S, Rau R, Steinfeld P. Very low-dose prednisolone in early rheumatoid arthritis retards radiographic progression over two years: A multicenter, double-blind, placebo-controlled trial. *Arthritis and Rheumatism.* 2005; 52(11): 3371-3380
  258. Kirwan JR, Bijlsma JWJ, BOERS M, Shea BJ. Effects of glucocorticoids on radiological progression in rheumatoid arthritis. *Cochrane Database Syst Rev.* 2007;(1):CD006356.
  259. Todoerti M, Scirè CA, Boffini N, Bugatti S, Montecucco C, Caporali R. Early disease control by low-dose prednisone comedication may affect the quality of remission in patients with early rheumatoid arthritis. *Annals of the New York Academy of Sciences.* 2010 Apr 13;1193(1):139–45.
  260. de Jong PH, Hazes JM, Barendregt PJ, Huisman M, van Zeben D, van der Lubbe PA, et al. Induction therapy with a combination of DMARDs is better than methotrexate monotherapy: first results of the tREACH trial. *Annals of the Rheumatic Diseases.* 2013 Jan;72(1):72–8.
  261. Chakravarty K, McDonald H, Pullar T, Taggart A, Chalmers R, Oliver S, et al. BSR/BHPR guideline for disease-modifying anti-rheumatic drug (DMARD) therapy in consultation with the British Association of Dermatologists. *Rheumatology (Oxford). Br Soc Rheumatology;* 2008;47(6):924–5.
  262. Wessels JAM, Huizinga TWJ, Guchelaar HJ. Recent insights in the pharmacological actions of methotrexate in the treatment of rheumatoid arthritis. *Rheumatology.* 2007 Nov 28;47(3):249–55.
  263. Cronstein BN, Montesinos MC, Weissmann G. Salicylates and sulfasalazine, but not glucocorticoids, inhibit leukocyte accumulation by an adenosine-dependent mechanism that is independent of inhibition of prostaglandin synthesis and p105 of NFκB. *Proc. Natl. Acad. Sci. U.S.A. National Acad Sciences;* 1999;96(11):6377–81.
  264. Morabito L, Montesinos MC, Schreiber DM, Balter L, Thompson LF, Resta R, et al. Methotrexate and sulfasalazine promote adenosine release by a mechanism that requires ecto-5'-nucleotidase-mediated conversion of adenine nucleotides. *J. Clin. Invest.* 1998 Jan 15;101(2):295–300.
  265. Volin MV, Campbell PL, Connors MA, Woodruff DC, Koch AE. The Effect of Sulfasalazine on Rheumatoid Arthritic Synovial Tissue Chemokine Production. *Experimental and Molecular Pathology.* 2002 Oct;73(2):84–92.
  266. Fox RI. Mechanism of action of hydroxychloroquine as an antirheumatic drug. *Seminars in Arthritis and Rheumatism. Elsevier;* 1993;23(2):82–91.
  267. The Royal College of Ophthalmologists Hydroxychloroquine and Ocular Toxicity Recommendations on Screening October 2009. The Royal College of

Ophthalmologists. 2009 Oct 8;:1–9.

268. Peppel K, Crawford D, Beutler B. A tumor necrosis factor (TNF) receptor-IgG heavy chain chimeric protein as a bivalent antagonist of TNF activity. *J. Exp. Med.* 1991 Dec 1;174(6):1483–9.
269. Choi HK, Hernán MA, Seeger JD, Robins JM, Wolfe F. Methotrexate and mortality in patients with rheumatoid arthritis: a prospective study. *The Lancet.* 2002 Apr;359(9313):1173–7.
270. O'Dell J, Haire C, Erikson N. Treatment of Rheumatoid Arthritis with Methotrexate Alone, Sulfasalazine and Hydroxychloroquine, or a Combination of All Three Medications. *NEJM.* 1996; 334: 1287-91
271. Calgüneri M, Pay S, Caliřkaner Z, Aprař S, Kiraz S, Ertenli I, et al. Combination therapy versus monotherapy for the treatment of patients with rheumatoid arthritis. *Clin. Exp. Rheumatol.* 1999;17(6):699.
272. Deighton C, Hyrich K, Ding T, Ledingham J, Lunt M, Luqmani R, et al. BSR and BHPR rheumatoid arthritis guidelines on eligibility criteria for the first biological therapy. *Rheumatology.* 2010 Mar 22.
273. Visser K, Katchamart W, Loza E, Martinez-Lopez JA, Salliot C, Trudeau J, et al. Multinational evidence-based recommendations for the use of methotrexate in rheumatic disorders with a focus on rheumatoid arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative. *Annals of the Rheumatic Diseases.* 2009 Jun 12;68(7):1086–93.
274. Herman RA, Veng Pedersen P, Hoffman J, Koehnke R, Furst DE. Pharmacokinetics of low-dose methotrexate in rheumatoid arthritis patients. *Journal of pharmaceutical sciences.* Wiley Online Library; 1989;78(2):165–71.
275. Jundt J, Browne B, Fiocco G, Steele A, Mock D. A comparison of low dose methotrexate bioavailability: oral solution, oral tablet, subcutaneous and intramuscular dosing. *The Journal of Rheumatology.* 1993;20(11):1845.
276. Hoekstra M, Haagsma C, Neef C, Proost J, Knuif A, van de Laar M. Bioavailability of higher dose methotrexate comparing oral and subcutaneous administration in patients with rheumatoid arthritis. *The Journal of Rheumatology.* *The Journal of Rheumatology;* 2004;31(4):645–8.
277. Rozin A, Schapira D, Balbir-Gurman A, Braun-Moscovici Y, Markovits D, Militianu D, et al. Relapse of rheumatoid arthritis after substitution of oral for parenteral administration of methotrexate. *Annals of the Rheumatic Diseases.* 2002 Aug;61(8):756–7.
278. Wegrzyn J, Adeleine P, Miossec P. Better efficacy of methotrexate given by intramuscular injection than orally in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 2004 Oct;63(10):1232–4.
279. Lambert CM, Sandhu S, Lochhead A, Hurst NP, McRorie E, Dhillon V. Dose escalation of parenteral methotrexate in active rheumatoid arthritis that has been unresponsive to conventional doses of methotrexate: a randomized, controlled trial. *Arthritis Rheum.* 2004 Feb;50(2):364–71.
280. Braun J, Kästner P, Flaxenberg P, Währisch J, Hanke P, Demary W, et al. Comparison of the clinical efficacy and safety of subcutaneous versus oral administration of methotrexate in patients with active rheumatoid arthritis: results of a six-month, multicenter, randomized, double-blind, controlled, phase IV trial. *Arthritis Rheum.* 2008 Jan;58(1):73–81.
281. Bathon JM, Martin RW, Fleischmann RM, Tesser JR, Schiff MH, Keystone EC, et

- al. A Comparison of Etanercept and Methotrexate in Patients with Early Rheumatoid Arthritis. *N Engl J Med*. 2000 Nov 30;343(22):1586–93.
282. Emery P, Breedveld FC, Hall S, Durez P, Chang DJ, Robertson D, et al. Comparison of methotrexate monotherapy with a combination of methotrexate and etanercept in active, early, moderate to severe rheumatoid arthritis (COMET): a randomised, double-blind, parallel treatment trial. *The Lancet*. 2008 Aug;372(9636):375–82.
283. Quinn MA, Conaghan PG, O'Connor PJ, Karim Z, Greenstein A, Brown A, et al. Very early treatment with infliximab in addition to methotrexate in early, poor-prognosis rheumatoid arthritis reduces magnetic resonance imaging evidence of synovitis and damage, with sustained benefit after infliximab withdrawal: Results from a twelve-month randomized, double-blind, placebo-controlled trial. *Arthritis Rheum*. 2005;52(1):27–35.
284. van der Bijl AE, Goekoop-Ruiterman YPM, de Vries-Bouwstra JK, Wolde ten S, Han KH, van Krugten MV, et al. Infliximab and methotrexate as induction therapy in patients with early rheumatoid arthritis. *Arthritis Rheum*. 2007;56(7):2129–34.
285. Remission Induction With Etanercept and Methotrexate In Very Early Rheumatoid Arthritis with Sustained Remission after Etanercept Withdrawal. *Arthritis Rheum*. 2008. p. S918.
286. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. 1990.
287. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, d'Agostino M-A, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *The Journal of Rheumatology*. 2005 Dec;32(12):2485–7.
288. Szkudlarek M, Court-Payen M, Jacobsen S, Klarlund M, Thomsen HS, ØStergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. *Arthritis Rheum*. 2003 Apr;48(4):955–62.
289. Hammer HB, Bolton-King P, Bakkeheim V, Berg TH, Sundt E, Kongtorp AK, et al. Examination of intra and interrater reliability with a new ultrasonographic reference atlas for scoring of synovitis in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 2011 Jul 22.
290. Scheel AK, Hermann K-GA, Kahler E, Pasewaldt D, Fritz J, Hamm B, et al. A novel ultrasonographic synovitis scoring system suitable for analyzing finger joint inflammation in rheumatoid arthritis. *Arthritis Rheum*. 2005;52(3):733–43.
291. Backhaus M, Ohrndorf S, Kellner H, Strunk J, Backhaus TM, Hartung W, et al. Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. *Arthritis Rheum*. 2009 Sep 15;61(9):1194–201.
292. Backhaus TM, Ohrndorf S, Kellner H, Strunk J, Hartung W, Sattler H, et al. The US7 score is sensitive to change in a large cohort of patients with rheumatoid arthritis over 12 months of therapy. *Annals of the Rheumatic Diseases*. 2012 Sep 6.
293. Dougados M, Jousse-Joulin S, Mistretta F, d'Agostino MA, Backhaus M, Bentin J, et al. Evaluation of several ultrasonography scoring systems for synovitis and comparison to clinical examination: results from a prospective multicentre study of rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 2010 Apr 22;69(5):828–33.
294. MA D, Wakefield R, Berner Hammer H, Vittecoq V, Galeazzi M, Balint P, et al. Assessment of Omeract Global Power Doppler Ultrasonography 44-Joint Scoring System and Reduced Joint Scoring Systems in Rheumatoid Arthritis Patients Treated with Abatacept Plus Background Methotrexate. *Arthritis Rheum*. 2012 Nov 18;64(10):S352.

295. Terslev L, Torp-Pedersen S, Qvistgaard E, Recke von der P, Bliddal H. Doppler ultrasound findings in healthy wrists and finger joints. *Annals of the Rheumatic Diseases*. 2004 Jun;63(6):644–8.
296. van der Heijde DM, van 't Hof M, van Riel PL, van De Putte LB. Development of a disease activity score based on judgment in clinical practice by rheumatologists. *The Journal of Rheumatology*. 1993 Mar;20(3):579–81.
297. van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van De Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum*. 1996 Jan;39(1):34–40.
298. Fries J, Spitz P, Kraines R. Measurement of patient outcome in arthritis - Fries - 2005 - *Arthritis & Rheumatism* - Wiley Online Library. *Arthritis & ...*. 1980.
299. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: Dimensions and Practical Applications. *Health Qual Life Outcomes*. BioMed Central Ltd; 2003;1(1):20.
300. Maska L, Anderson J, Michaud K. Measures of functional status and quality of life in rheumatoid arthritis: Health Assessment Questionnaire Disability Index (HAQ), Modified Health Assessment Questionnaire (MHAQ), Multidimensional Health Assessment Questionnaire (MDHAQ), Health Assessment. *Arthritis Care Res*. 2011 Nov 7;63(S11):S4–S13.
301. EQ-5D-3L User Guide. 2012 Nov 20;:1–24.
302. Felson DT, Anderson JJ, BOERS M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis Rheum*. Wiley Online Library; 1993;36(6):729–40.
303. OSTERGAARD M. An introduction to the EULAR-OMERACT rheumatoid arthritis MRI reference image atlas. *Annals of the Rheumatic Diseases*. 2005 Feb 1;64(suppl\_1):i3–i7.
304. Conaghan P. The EULAR-OMERACT rheumatoid arthritis MRI reference image atlas: the metacarpophalangeal joints. *Annals of the Rheumatic Diseases*. 2005 Feb 1;64(suppl\_1):i11–i21.
305. Ejbjerg B. The EULAR-OMERACT rheumatoid arthritis MRI reference image atlas: the wrist joint. *Annals of the Rheumatic Diseases*. 2005 Feb 1;64(suppl\_1):i23–i47.
306. McQueen FM, Stewart N, Crabbe J, Robinson E, Yeoman S, Tan PL, et al. Magnetic resonance imaging of the wrist in early rheumatoid arthritis reveals progression of erosions despite clinical improvement. *Annals of the Rheumatic Diseases*. 1999 Mar;58(3):156–63.
307. Ejbjerg BJ, Vestergaard A, Jacobsen S, Thomsen HS, ØStergaard M. The smallest detectable difference and sensitivity to change of magnetic resonance imaging and radiographic scoring of structural joint damage in rheumatoid arthritis finger, wrist, and toe joints: A comparison of the omeract rheumatoid arthritis magnetic resonance imaging score applied to different joint combinations and the sharp/van der heijde radiographic score. *Arthritis Rheum*. 2005;52(8):2300–6.
308. Tan YK, Connaghan P. Imaging in rheumatoid arthritis. *Best Practice and Research Clinical Rheumatology*. 2011 Jun 1;25:569–84.
309. van der Heijde DM, van Leeuwen MA, van Riel PL, Koster AM, van 't Hof MA, van Rijswijk MH, et al. Biannual radiographic assessments of hands and feet in a three-year prospective followup of patients with early rheumatoid arthritis. *Arthritis*

- Rheum. 1992 Jan;35(1):26–34.
310. Baron G, Boutron I, Giraudeau B, Ravaud P. Reporting of radiographic methods in randomised controlled trials assessing structural outcomes in rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 2007 May 1;66(5):651–7.
  311. van der Heijde D, Dankert T, Nieman F, Rau R, BOERS M. Reliability and sensitivity to change of a simplification of the Sharp/van der Heijde radiological assessment in rheumatoid arthritis. *Rheumatology (Oxford)*. 1999 Oct;38(10):941–7.
  312. Boini S, Guillemin F. Radiographic scoring methods as outcome measures in rheumatoid arthritis: properties and advantages. *Annals of the Rheumatic Diseases*. 2001 Sep;60(9):817–27.
  313. van der Heijde D, Boonen A, BOERS M, Kostense P, van der Linden S. Reading radiographs in chronological order, in pairs or as single films has important implications for the discriminative power of rheumatoid arthritis clinical trials. *Rheumatology (Oxford)*. *Br Soc Rheumatology*; 1999;38(12):1213–20.
  314. ALETAHA D, Landewe R, Karonitsch T, Bathon J, BOERS M, Bombardier C, et al. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. *Arthritis Rheum*. 2008 Oct 15;59(10):1371–7.
  315. Felson DT, Anderson JJ, BOERS M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum*. 1995 Jun;38(6):727–35.
  316. Van Gelder RN, Zastrow von ME, Yool A, Dement WC, Barchas JD, Eberwine JH. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 1990 Mar;87(5):1663–7.
  317. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet*. 2007 Sep;3(9):1724–35.
  318. Smyth GK. Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Statistical Applications in Genetics and Molecular Biology*. 3(1).
  319. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*. JSTOR; 1995;:289–300.
  320. Goetz I, Carter GC, Lucero M, Zarotsky V, Alatorre CI, Cantrell RA, et al. Review of treatment response in rheumatoid arthritis: assessment of heterogeneity. *Curr Med Res Opin*. 2011 Apr;27(4):697–711.
  321. Krabben A, Huizinga TWJ, van der Helm-van Mil AHM. Undifferentiated arthritis characteristics and outcomes when applying the 2010 and 1987 criteria for rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 2011 Dec 28;71(2):238–41.
  322. Snir O, Widhe M, Spee von C, Lindberg J, Padyukov L, Lundberg K, et al. Multiple antibody reactivities to citrullinated antigens in sera from patients with rheumatoid arthritis: association with HLA-DRB1 alleles. *Annals of the Rheumatic Diseases*. 2009 May 1;68(5):736–43.
  323. Goeldner I, Skare TL, de Messias Reason IT, Nishihara RM, Silva MB, da Rosa Utiyama SR. Anti-cyclic citrullinated peptide antibodies and rheumatoid factor in rheumatoid arthritis patients and relatives from Brazil. *Rheumatology*. 2010 Jul 13;49(8):1590–3.
  324. Agrawal S, Misra R, Aggarwal A. Autoantibodies in rheumatoid arthritis: association

with severity of disease in established RA. *Clin. Rheumatol.* 2006 Mar 30;26(2):201–4.

325. Fautrel B, Combe B, Rincheval N, Dougados M, ESPOIR Scientific Committee. Level of agreement of the 1987 ACR and 2010 ACR/EULAR rheumatoid arthritis classification criteria: an analysis based on ESPOIR cohort data. *Annals of the Rheumatic Diseases.* 2012 Mar;71(3):386–9.
326. Nicolau J, Boumier P, Daragon A, Mejjad O, Menard J-F, Pouplin S, et al. VALIDATION OF THE 2010 ACR/EULAR CLASSIFICATION CRITERIA FOR RHEUMATOID ARTHRITIS IN THE VERY EARLY ARTHRITIS COMMUNITY-BASED (VERA) COHORT: SLIGHT IMPROVEMENT AFTER EXCLUDING EROSION CRITERION. *Annals of the Rheumatic Diseases.* 2012. p. 98.
327. Radner H, Neogi T, Smolen JS, ALETAHA D. Performance of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: a systematic literature review. *Annals of the Rheumatic Diseases.* 2013 Apr 16.
328. Luchikhina EL, Karateev DE, Novikov AA, Alexandrova EN, Lukina GV, Nasonov EL. PERFORMANCE AND PREDICTIVE VALUE OF ACR/EULAR 2010 RHEUMATOID ARTHRITIS CLASSIFICATION CRITERIA IN DIFFERENT GROUPS OF PATIENTS WITH EARLY ARTHRITIS. *Annals of the Rheumatic Diseases.* 2011. p. 285.
329. Wells George A, Boers M, Shea B, Brooks PM, Simon LS, Strand CV, et al. Minimal disease activity for rheumatoid arthritis: a preliminary definition. *J. Rheumatol.* 2005. pp. 2016–24.
330. Brown AK. Using ultrasonography to facilitate best practice in diagnosis and management of RA. *Nature Reviews Rheumatology.* Nature Publishing Group; 2009 Nov 10;5(12):698–706.
331. Mandl P, Balint PV, Brault Y, Backhaus M, d'Agostino MA, Grassi W, et al. Metrologic properties of ultrasound versus clinical evaluation of synovitis in rheumatoid arthritis: Results of a multicenter, randomized study. *Arthritis Rheum.* 2012 Mar 27;64(4):1272–82.
332. Naredo E, Collado P, Cruz A, Palop MJ, Cabero F, Richi P, et al. Longitudinal power Doppler ultrasonographic assessment of joint inflammatory activity in early rheumatoid arthritis: Predictive value in disease activity and radiologic progression. *Arthritis Care and Research.* 2007; 57(1): 116–124
333. Qvistgaard E, Røgind H, Torp-Pedersen S, Terslev L, Danneskiold-Samsøe B, Bliddal H. Quantitative ultrasonography in rheumatoid arthritis: evaluation of inflammation by Doppler technique. *Annals of the Rheumatic Diseases.* 2001 Jul;60(7):690–3.
334. Carotti M, Salaffi F, Morbiducci J, Ciapetti A, Bartolucci L, Gasparini S, et al. Colour Doppler ultrasonography evaluation of vascularization in the wrist and finger joints in rheumatoid arthritis patients and healthy subjects. *European Journal of Radiology.* 2012 Aug;81(8):1834–8.
335. Commission A. Outpatients. Audit Commission. 2003 Jun 1;:1–23.
336. Saleem B, Brown AK, Keen H, Nizam S, Freeston J, Karim Z, et al. Disease remission state in patients treated with the combination of tumor necrosis factor blockade and methotrexate or with disease-modifying antirheumatic drugs: A clinical and imaging comparative study. *Arthritis Rheum.* 2009 Jul;60(7):1915–22.
337. Saleem B, Brown AK, Keen H, Nizam S, Freeston J, Wakefield R, et al. Should imaging be a component of rheumatoid arthritis remission criteria? A comparison between traditional and modified composite remission scores and imaging assessments. *Annals of the Rheumatic Diseases.* 2011 Apr 3;70(5):792–8.

338. Listing J, Gerhold K, Zink A. The risk of infections associated with rheumatoid arthritis, with its comorbidity and treatment. *Rheumatology*. 2012 Dec 13;52(1):53–61.
339. Saleem B, Brown AK, Quinn M, Karim Z, Hensor EMA, Conaghan P, et al. Can flare be predicted in DMARD treated RA patients in remission, and is it important? A cohort study. *Annals of the Rheumatic Diseases*. 2012 Jul 14;71(8):1316–21.
340. Staud R. Brain imaging in fibromyalgia syndrome. *Clin. Exp. Rheumatol*. 2011 Nov;29(6 Suppl 69):S109–17.
341. Marchesoni A, DE LUCIA O, ROTUNNO L, DE MARCO G, MANARA M. Enthesial Power Doppler Ultrasonography: A Comparison of Psoriatic Arthritis and Fibromyalgia. *The Journal of Rheumatology Supplement*. 2012 Jul 1;89(0):29–31.
342. Keen HI, Lavie F, Wakefield RJ, d'Agostino MA, Hammer HB, Hensor E, et al. The development of a preliminary ultrasonographic scoring system for features of hand osteoarthritis. *Annals of the Rheumatic Diseases*. 2007 Aug 17;67(5):651–5.
343. Iagnocco A. Imaging the joint in osteoarthritis: a place for ultrasound? *Best Pract Res Clin Rheumatol*. 2010 Feb;24(1):27–38.
344. National Collaborating Centre for Chronic Conditions (UK). Osteoarthritis: National clinical guideline for care and management in adults. London: Royal College of Physicians (UK); 2008.
345. Dziedzic K, Thomas E, Hill S, Wilkie R, Peat G, Croft PR. The impact of musculoskeletal hand problems in older adults: findings from the North Staffordshire Osteoarthritis Project (NorStOP). *Rheumatology*. 2007 Mar 27;46(6):963–7.
346. Wenham CYJ, Grainger AJ, Hensor EMA, Caperon AR, Ash ZR, Conaghan PG. Methotrexate for pain relief in knee osteoarthritis: an open-label study. *Rheumatology*. 2013 Apr 19;52(5):888–92.
347. Bryant LR, Rosier des KF, Carpenter MT. Hydroxychloroquine in the treatment of erosive osteoarthritis. *The Journal of Rheumatology*. 1995 Aug;22(8):1527–31.
348. Guler-Yuksel M, Allaart CF, Watt I, Goekoop-Ruiterman YPM, de Vries-Bouwstra JK, van Schaardenburg D, et al. Treatment with TNF- $\alpha$  inhibitor infliximab might reduce hand osteoarthritis in patients with rheumatoid arthritis. *Osteoarthritis and Cartilage*. 2010 Oct;18(10):1256–62.
349. Keen HI, Wakefield RJ, Hensor EMA, Emery P, Conaghan PG. Response of symptoms and synovitis to intra-muscular methylprednisolone in osteoarthritis of the hand: an ultrasonographic study. *Rheumatology*. 2010 May 13;49(6):1093–100.
350. Schoels M, Knevel R, ALETAHA D, Bijlsma JWJ, Breedveld FC, Boumpas DT, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search. *Annals of the Rheumatic Diseases*. 2010 Mar 17;69(4):638–43.
351. Naredo E, Möller I, Cruz A, Carmona L, Garrido J. Power doppler ultrasonographic monitoring of response to anti-tumor necrosis factor therapy in patients with rheumatoid arthritis. *Arthritis Rheum*. 2008 Aug;58(8):2248–56.
352. Naredo E, Collado P, Cruz A, Palop MJ, Cabero F, Richi P, et al. Longitudinal power Doppler ultrasonographic assessment of joint inflammatory activity in early rheumatoid arthritis: Predictive value in disease activity and radiologic progression. *Arthritis Rheum*. 2007;57(1):116–24.
353. Beineke P, Fitch K, Tao H, Elashoff MR, Rosenberg S, Kraus WE, et al. A whole blood gene expression-based signature for smoking status. *BMC Med Genomics*. 2012;5:58.

354. Buro-Auriemma LJ, Salit J, Hackett NR, Walters MS, Strulovici-Barel Y, Staudt MR, et al. Cigarette Smoking Induces Small Airway Epithelial Epigenetic Changes with Corresponding Modulation of Gene Expression. *Hum. Mol. Genet.* 2013 Jul 10.
355. Raterman HG, Vosslander S, de Ridder S, Nurmohamed MT, Lems WF, Boers M, et al. The interferon type I signature towards prediction of non-response to rituximab in rheumatoid arthritis patients. *Arthritis Res Ther.* 2012;14(2):R95.
356. Vosslander S, Raterman HG, van der Pouw Kraan TCTM, Schreurs MWJ, Blomberg von BME, Nurmohamed MT, et al. Pharmacological induction of interferon type I activity following treatment with rituximab determines clinical response in rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 2011 Jun;70(6):1153–9.
357. Tabit CE, Holbrook M, Shenouda SM, Dohadwala MM, Widlansky ME, Frame AA, et al. Effect of sulfasalazine on inflammation and endothelial function in patients with established coronary artery disease. *Vascular Medicine.* 2012 Apr 10;17(2):101–7.
358. Whitney A, Diehn M, Popper S, Alizadeh AA, Boldrick J, Relman D, et al. Individuality and variation in gene expression patterns in human blood. *PNAS.* 2003 Feb 18;100(4):1896–901.
359. Liberman AC, Druker J, Perone MJ, Arzt E. Glucocorticoids in the regulation of transcription factors that control cytokine synthesis. *Cytokine & Growth Factor Reviews.* 2007 Feb;18(1-2):45–56.
360. Karlovich C, Duchateau-Nguyen G, Johnson A, McLoughlin P, Navarro M, Fleurbaey C, et al. A longitudinal study of gene expression in healthy individuals. *BMC Med Genomics.* 2009;2:33.
361. Zen M, Canova M, Campana C, Bettio S, Nalotto L, Rampudda M, et al. The kaleidoscope of glucocorticoid effects on immune system. *Autoimmunity Reviews.* 2011 Apr;10(6):305–10.
362. Xu Q, Ni S, Wu F, Liu F, Ye X, Mouglin B, et al. Investigation of Variation in Gene Expression Profiling of Human Blood by Extended Principle Component Analysis. Emanuelli C, editor. *PLoS ONE.* 2011 Oct 27;6(10):e26905.
363. Charlesworth JC, Curran JE, Johnson MP, Göring HH, Dyer TD, Diego VP, et al. Transcriptomic epidemiology of smoking: the effect of smoking on gene expression in lymphocytes. *BMC Med Genomics.* 2010;3:29.
364. Sugiyama D, Nishimura K, Tamaki K, Tsuji G, Nakazawa T, Morinobu A, et al. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. *Annals of the Rheumatic Diseases.* 2010 Jan;69(1):70–81.
365. Walker DJ, Pound JD, Griffiths ID, Powell RJ. Rheumatoid factor tests in the diagnosis and prediction of rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 1986 Aug;45(8):684–90.
366. Shmerling RH. How Useful Is the Rheumatoid Factor? An Analysis of Sensitivity, Specificity, and Predictive Value. *Arch Intern Med.* 1992 Dec 1;152(12):2417.
367. Bovin LF, Rieneck K, Workman C, Nielsen H, Sørensen SF, Skjødt H, et al. Blood cell gene expression profiling in rheumatoid arthritis. *Immunology Letters.* 2004 May;93(2-3):217–26.
368. van Baarsen LGM, Bos WH, Rustenburg F, van der Pouw Kraan TCTM, Wolbink GJJ, Dijkmans BAC, et al. Gene expression profiling in autoantibody-positive patients with arthralgia predicts development of arthritis. *Arthritis Rheum.* 2010 Feb 25;62(3):694–704.

369. Cantaert T, Brouard S, Thurlings RM, Pallier A, Salinas GF, Braud C, et al. Alterations of the synovial T cell repertoire in anti-citrullinated protein antibody-positive rheumatoid arthritis. *Arthritis Rheum.* 2009 Jul;60(7):1944–56.
370. Junta CM, Sandrin-Garcia P, Fachin-Saltoratto AL, Mello SS, Oliveira RDR, Rassi DM, et al. Differential gene expression of peripheral blood mononuclear cells from rheumatoid arthritis patients may discriminate immunogenetic, pathogenic and treatment features. *Immunology.* 2009 Jul;127(3):365–72.
371. van Baarsen LGM, Wijbrandts CA, Timmer TCG, van der Pouw Kraan TCTM, Tak PP, Verweij CL. Synovial tissue heterogeneity in rheumatoid arthritis in relation to disease activity and biomarkers in peripheral blood. *Arthritis Rheum.* 2010 Feb 22;62(6):1602–7.
372. Reynolds RJ, Cui X, Vaughan LK, Redden DT, Causey Z, Perkins E, et al. Gene expression patterns in peripheral blood cells associated with radiographic severity in African Americans with early rheumatoid arthritis. *Rheumatol Int.* 2012 Jan 12.
373. Tak PP, Bresnihan B. The pathogenesis and prevention of joint damage in rheumatoid arthritis: advances from synovial biopsy and tissue analysis. *Arthritis Rheum.* 2000 Dec;43(12):2619–33.
374. Hider SL, Bruce IN, Thomson W. The pharmacogenetics of methotrexate. *Rheumatology.* 2007 Aug 27;46(10):1520–4.
375. OLIVEIRA RDR, FONTANA V, JUNTA CM, MARQUES MMC, MACEDO C, RASSI DM, et al. Differential Gene Expression Profiles May Differentiate Responder and Nonresponder Patients with Rheumatoid Arthritis for Methotrexate (MTX) Monotherapy and MTX plus Tumor Necrosis Factor Inhibitor Combined Therapy. *The Journal of Rheumatology.* 2012 Aug 1;39(8):1524–32.
376. Meugnier E, Coury F, Tebib J, Ferraro-Peyret C, Rome S, Bienvenu J, et al. Gene expression profiling in peripheral blood cells of patients with rheumatoid arthritis in response to anti-TNF-alpha treatments. *Physiol. Genomics.* 2011 Apr 12;43(7):365–71.
377. Galligan CL, Baig E, Bykerk V, Keystone EC, Fish EN. Distinctive gene expression signatures in rheumatoid arthritis synovial tissue fibroblast cells: correlates with disease activity. *Genes Immun.* 2007 Jun 14;8(6):480–91.
378. Edwards C. Combined anti-tumor necrosis factor- $\alpha$  therapy and DMARD therapy in rheumatoid arthritis patients reduces inflammatory gene expression in whole blood compared to DMARD therapy alone. *Front. Immunol.* 2012;3.
379. Julià A, Erra A, Palacio C, Tomas C, Sans X, Barceló P, et al. An eight-gene blood expression profile predicts the response to infliximab in rheumatoid arthritis. *PLoS ONE.* 2009;4(10):e7556.
380. Badot V, Galant C, Nzeusseu Toukap A, Theate I, Maudoux A-L, Van den Eynde BJ, et al. Gene expression profiling in the synovium identifies a predictive signature of absence of response to adalimumab therapy in rheumatoid arthritis. *Arthritis Res Ther.* 2009;11(2):R57.
381. Thurlings RM, Boumans M, Tekstra J, van Roon JA, Vos K, van Westing DM, et al. Relationship between the type I interferon signature and the response to rituximab in rheumatoid arthritis patients. *Arthritis Rheum.* 2010 Dec;62(12):3607–14.
382. Bansard C, Lequerré T, Derambure C, Vittecoq O, Hiron M, Daragon A, et al. Gene profiling predicts rheumatoid arthritis responsiveness to IL-1Ra (anakinra). *Rheumatology.* 2011 Feb;50(2):283–92.
383. Centola M, Cavet G, Shen Y, Ramanujan S, Knowlton N, Swan KA, et al. Development of a Multi-Biomarker Disease Activity Test for Rheumatoid Arthritis.

Frey O, editor. PLoS ONE. 2013 Apr 9;8(4):e60635.

384. van Baarsen LG, Wijbrandts CA, Rustenburg F, Cantaert T, van der Pouw Kraan TC, Baeten DL, et al. Regulation of IFN response gene activity during infliximab treatment in rheumatoid arthritis is associated with clinical response to treatment. *Arthritis Res Ther.* 2010;12(1):R11.
385. van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Rustenburg F, Baggen JM, Verweij CL, et al. Responsiveness to anti-tumour necrosis factor therapy is related to pre-treatment tissue inflammation levels in rheumatoid arthritis patients. *Annals of the Rheumatic Diseases.* 2007 Aug 2;67(4):563–6.
386. Lubbers J, Brink M, van de Stadt LA, Vosslamber S, Wesseling JG, van Schaardenburg D, et al. The type I IFN signature as a biomarker of preclinical rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 2013 Apr 4;72(5):776–80.
387. Tan FK, Zhou X, Mayes MD, Gourh P, Guo X, Marcum C, et al. Signatures of differentially regulated interferon gene expression and vasculotrophism in the peripheral blood cells of systemic sclerosis patients. *Rheumatology (Oxford).* 2006 Jun;45(6):694–702.
388. Allantaz F, Chaussabel D, Stichweh D, Bennett L, Allman W, Mejias A, et al. Blood leukocyte microarrays to diagnose systemic onset juvenile idiopathic arthritis and follow the response to IL-1 blockade. *J. Exp. Med.* 2007 Sep 3;204(9):2131–44.
389. Maas K, Chen H, Shyr Y, Olsen NJ, Aune T. Shared gene expression profiles in individuals with autoimmune disease and unaffected first-degree relatives of individuals with autoimmune disease. *Hum. Mol. Genet.* 2005 May 15;14(10):1305–14.
390. Curtis JR, van der Helm-van Mil AH, Knevel R, Huizinga TW, Haney DJ, Shen Y, et al. Validation of a novel multibiomarker test to assess rheumatoid arthritis disease activity. *Arthritis Care Res.* 2012 Nov 28;64(12):1794–803.
391. Taylor PC, Steuer A, Gruber J, McClinton C, Cosgrove DO, Blomley MJK, et al. Ultrasonographic and radiographic results from a two-year controlled trial of immediate or one-year–delayed addition of infliximab to ongoing methotrexate therapy in patients with erosive early rheumatoid arthritis. *Arthritis Rheum.* 2005;54(1):47–53.
392. Bakker MF, Cavet G, Jacobs JW, Bijlsma JWJ, Haney DJ, Shen Y, et al. Performance of a multi-biomarker score measuring rheumatoid arthritis disease activity in the CAMERA tight control study. *Annals of the Rheumatic Diseases.* 2012 Sep 10;71(10):1692–7.
393. Hirata S, Dirven L, Shen Y, Centola M, Cavet G, Lems WF, et al. A multi-biomarker score measures rheumatoid arthritis disease activity in the BeSt study. *Rheumatology.* 2013 Jun 18;52(7):1202–7.
394. van der Helm-van Mil AHM, Knevel R, Cavet G, Huizinga TWJ, Haney DJ. An evaluation of molecular and clinical remission in rheumatoid arthritis by assessing radiographic progression. *Rheumatology.* 2013 Apr 19;52(5):839–46.
395. Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature.* 2012 Apr 18.
396. Bakker MF, Jacobs JW, Welsing PMJ, Vreugdenhil SA, van Booma-Frankfort C, Linn-Rasker SP, et al. Early clinical response to treatment predicts 5-year outcome in RA patients: follow-up results from the CAMERA study. *Annals of the Rheumatic Diseases.* 2011 May 2;70(6):1099–103.
397. Peabody JW, Strand V, Shimkhada R, Lee R, Chernoff D. Impact of Rheumatoid

Arthritis Disease Activity Test on Clinical Practice. Frey O, editor. PLoS ONE. 2013 May 7;8(5):e63215.

398. Li W, Sasso EH, Emerling D, Cavet G, Ford K. Impact of a multi-biomarker disease activity test on rheumatoid arthritis treatment decisions and therapy use. *Curr Med Res Opin.* 2013 Jan;29(1):85–92.
399. Hambardzumyan K, Bolce R, Cavet G, Chernoff D, Haney DJ, Forslid K, et al. A MULTI-BIOMARKER DISEASE ACTIVITY BLOOD TEST (VECTRA DA) CORRELATES WITH RADIOGRAPHIC PROGRESSION IN EARLY RHEUMATOID ARTHRITIS: RESULTS FROM THE SWEFOT TRIAL. *Annals of the Rheumatic Diseases.* 2013. p. 388.
400. Szkudlarek M, Court-Payen M, Strandberg C, Klarlund M, Klausen T, OSTERGAARD M. Power Doppler ultrasonography for assessment of synovitis in the metacarpophalangeal joints of patients with rheumatoid arthritis: a comparison with dynamic magnetic resonance imaging. *Arthritis Rheum.* 2001 Sep;44(9):2018–23.
401. Naredo E. Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with grey scale and power Doppler ultrasonography. *Annals of the Rheumatic Diseases.* 2004 Jul 22;64(3):375–81.
402. Reuther LO, Vainer B, Sonne J, Larsen N-E. Thiopurine methyltransferase (TPMT) genotype distribution in azathioprine-tolerant and -intolerant patients with various disorders. The impact of TPMT genotyping in predicting toxicity. *Eur. J. Clin. Pharmacol.* 2004 Jan;59(11):797–801.
403. Smith SL, Plant D, Eyre S, Barton A. The potential use of expression profiling: implications for predicting treatment response in rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 2013 May 30;72(7):1118–24.
404. van Baarsen LG, Bos CL, Pouw Kraan T, Verweij CL. Transcription profiling of rheumatic diseases. *Arthritis Res Ther.* 2009;11(1):207.
405. Haagsma CJ, van Riel PLCM, DE ROOIJ DJRAM, VREE TB, RUSSEL FJM, van't Hof MA, et al. COMBINATION OF METHOTREXATE AND SULPHASALAZINE VS METHOTREXATE ALONE: A RANDOMIZED OPEN CLINICAL TRIAL IN RHEUMATOID ARTHRITIS PATIENTS RESISTANT TO SULPHASALAZINE THERAPY. *Rheumatology.* 1994;33(11):1049–55.
406. Haavardsholm EA, Boyesen P, OSTERGAARD M, Schildvold A, Kvien TK. Magnetic resonance imaging findings in 84 patients with early rheumatoid arthritis: bone marrow oedema predicts erosive progression. *Annals of the Rheumatic Diseases.* 2008 Jan 21;67(6):794–800.
407. van Vollenhoven R, Ernestam S, Geborek P, Petersson I, Coster L et al. Addition of infliximab compared with addition of sulfasalazine and hydroxychloroquine to methotrexate in patients with early rheumatoid arthritis (Swefot trial): 1-year results of a randomised trial. *The Lancet.* 2009; 374: 459-466
408. O'Dell JR, Mikuls TR, Taylor TH, Ahluwalia V, Brophy M, Warren SR, et al. Therapies for Active Rheumatoid Arthritis after Methotrexate Failure. *N Engl J Med.* 2013 Jun 11;:130611150115007.
409. Ziswiler HR, Aeberli D, Villiger PM, Moller B. High-resolution ultrasound confirms reduced synovial hyperplasia following rituximab treatment in rheumatoid arthritis. *Rheumatology.* 2009 Jul 12;48(8):939–43.
410. Albrecht K, Grob K, Lange U, Muller-Ladner U, Strunk J. Reliability of different Doppler ultrasound quantification methods and devices in the assessment of therapeutic response in arthritis. *Rheumatology.* 2008 Aug 13;47(10):1521–6.

411. Filippucci E. Grey scale and power Doppler sonographic changes induced by intra-articular steroid injection treatment. *Annals of the Rheumatic Diseases*. 2004 Mar 17;63(6):740–3.
412. Filippucci E, Iagnocco A, Salaffi F, Cerioni A, Valesini G, Grassi W. Power Doppler sonography monitoring of synovial perfusion at the wrist joints in patients with rheumatoid arthritis treated with adalimumab. *Annals of the Rheumatic Diseases*. 2006 Nov 1;65(11):1433–7.
413. Haavardsholm EA, OSTERGAARD M, Hammer HB, Boyesen P, Boonen A, van der Heijde D, et al. Monitoring anti-TNF treatment in rheumatoid arthritis: responsiveness of magnetic resonance imaging and ultrasonography of the dominant wrist joint compared with conventional measures of disease activity and structural damage. *Annals of the Rheumatic Diseases*. 2009 Sep 11;68(10):1572–9.
414. Trouw LA, Huizinga TWJ, Toes REM. Autoimmunity in rheumatoid arthritis: different antigens--common principles. *Annals of the Rheumatic Diseases*. 2013 Mar 25;72(Supplement 2):ii132–6.
415. Julià A, Erra A, Palacio C, Tomas C, Sans X, Barceló P, et al. An Eight-Gene Blood Expression Profile Predicts the Response to Infliximab in Rheumatoid Arthritis. Hoheisel J, editor. *PLoS ONE*. 2009 Oct 22;4(10):e7556.
416. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. *The Lancet*. 2011 Nov;378(9805):1812–23.
417. Taurino C, Miller WH, McBride MW, McClure JD, Khanin R, Moreno MU, et al. Gene expression profiling in whole blood of patients with coronary artery disease. *Clin. Sci*. 2010 Jul 9;119(8):335–43.
418. Zhao C, Mao J, Ai J, Shenwu M, Shi T, Zhang D, et al. Integrated lipidomics and transcriptomic analysis of peripheral blood reveals significantly enriched pathways in type 2 diabetes mellitus. *BMC Med Genomics*. 2013;6 Suppl 1:S12.
419. Ballantine LE. Transcriptomic analysis of peripheral blood monocytes and synovial macrophages in inflammatory arthritis. University of Glasgow; 2011.
420. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, et al. Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet*. 2010 Oct;11(10):733–9.
421. Debey S, Schoenbeck U, Hellmich M, Gathof BS, Pillai R, Zander T, et al. Comparison of different isolation techniques prior gene expression profiling of blood derived cells: impact on physiological responses, on overall expression and the role of different cell types. *Pharmacogenomics J*. 2004;4(3):193–207.
422. Asare AL, Kolchinsky SA, Gao Z, Wang R, Raddassi K, Bourcier K, et al. Differential gene expression profiles are dependent upon method of peripheral blood collection and RNA isolation. *BMC Genomics*. 2008;9(1):474.
423. Menke A, Rex-Haffner M, Klengel T, Binder EB, Mehta D. Peripheral blood gene expression: it all boils down to the RNA collection tubes. *BMC Res Notes*. 2012;5:1.
424. Kennedy L, Vass JK, Haggart DR, Moore S, Burczynski ME, Crowther D, et al. Hematopoietic Lineage Transcriptome Stability and Representation in PAXgene™ Collected Peripheral Blood Utilising SPIA Single-Stranded cDNA Probes for Microarray. *Biomarker insights*. *Libertas Academica*; 2008;3:403.
425. Min JL, Barrett A, Watts T, Pettersson FH, Lockstone HE, Lindgren CM, et al. Variability of gene expression profiles in human blood and lymphoblastoid cell lines. *BMC Genomics*. 2010;11:96.
426. Liu J, Walter E, Stenger D, Thach D. Effects of Globin mRNA Reduction Methods

- on Gene Expression Profiles from Whole Blood. *The Journal of Molecular Diagnostics*. 2006 Nov;8(5):551–8.
427. Tan PK, Downey TJ, Spitznagel EL, Xu P, Fu D, Dimitrov DS, et al. Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res*. 2003 Oct 1;31(19):5676–84.
428. Weis BK. Standardizing global gene expression analysis between laboratories and across platforms. *Nat Meth*. 2005 Apr 21;2(5):351–6.
429. Yauk CL, Berndt ML. Review of the literature examining the correlation among DNA microarray technologies. *Environ. Mol. Mutagen*. 2007;48(5):380–94.
430. Pan W. A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments. *Bioinformatics*. 2002 Apr;18(4):546–54.
431. Jeffery IB, Higgins DG, Culhane AC. Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data. *BMC Bioinformatics*. 2006;7(1):359.
432. Jeanmougin M, de Reynies A, Marisa L, Paccard C, Nuel G, Guedj M. Should We Abandon the t-Test in the Analysis of Gene Expression Microarray Data: A Comparison of Variance Modeling Strategies. Shedden K, editor. *PLoS ONE*. 2010 Sep 3;5(9):e12336.
433. Bansard C, Lequerre T, Daveau M, Boyer O, Tron F, Salier JP, et al. Can rheumatoid arthritis responsiveness to methotrexate and biologics be predicted? *Rheumatology*. 2009 Aug 17;48(9):1021–8.

## Appendix A - Abbreviations

ACR	American College of Rheumatology
ACPA	Anti-Citrillunated Protein Antibodies
ANA	Anti-nuclear antibody
AUC	Area Under the Curve
CCP	Cyclic-citrillunated peptide
CD	Colour Doppler MSUS Examination
CDAI	Clinical Disease Activity Index
COMB	Combination DMARD Therapy
CRP	C-Reactive Protein
DAS28	28 joint Disease Activity Score
DAS44	44 joint Disease Activity Score
DMARD	Disease Modifying Anti-Rheumatic Drug
ESR	Erythrocyte sedimentation rate
EULAR	European League Against Rheumatism
FBC	Full blood count
FDA	Federal Drugs Agency
HAQ	Health Assessment Questionnaire
HAQ-DI	Health Assessment Questionnaire Disability Index
IRG	Interferon response genes
LDAS	Low disease activity score
LFT	Liver function tests
MBDA	Multibiomarker Disease Activity Test
MCPj	Metacarpophalangeal Joint
MRI	Magnetic Resonance Imaging
MSUS	Musculoskeletal Ultrasound
MTPj	Metatarsophalangeal Joint
MTX	Methotrexate
NHSGGC	NHS Greater Glasgow and Clyde
OMERACT	Outcome Measures in Rheumatology
OA	Osteoarthritis
PD	Power Doppler MSUS Examination
PAS / PASII	Patient Activity Scores
PBMC	Peripheral blood mononuclear cell
PiPj	Proximal Interphalangeal Joint
RA	Rheumatoid Arthritis
RADAI	Rheumatoid Arthritis Disease Activity Index
RAI	Ritchie Articular Index
RAPID	Routine Assessment of Patient Index Data
RF	Rheumatoid Factor
ROC	Receiver Operating Characteristics
SASP	Sulfasalazine
SERA	Scottish Early Rheumatoid Arthritis Inception Cohort
SDAI	Simplified Disease Activity Index
SNP	Single nucleotide polymorphism
TFT	Thyroid function test
TMRC	Translational Medicine Research Collaboration
UE	Urea and Electrolytes
UIA	Undifferentiated Inflammatory Arthritis
ULN	Upper limit of normal

## Appendix B – Recruitment Poster



University  
of Glasgow



# TaSER

## Targeting Synovitis in Early Rheumatoid Arthritis

<b>Objective</b>	to compare the effectiveness of using musculoskeletal ultrasound or DAS28 to assess global disease activity and guide DMARD escalation in early rheumatoid arthritis using a tight control regimen
<b>Sites</b>	Gartnavel General Hospital, Glasgow Royal Infirmary, Stobhill Hospital
<b>Patient Involvement</b>	Monthly reviews at local early arthritis clinic for 18 months.

### Inclusion Criteria

1. Patients attending Early RA clinics with newly diagnosed RA or anti-CCP +ve Undifferentiated Arthritis (with 3 or more swollen joints)
2. Symptom duration <12 months
3. Active disease (DAS44>2.4)
4. DMARD naïve or DMARD monotherapy for less than 6 weeks
5. Aged 18 or more

### Exclusion Criteria

1. significant liver disease and/or abnormality of liver function tests  
- baseline AST / ALT > x2 normal, Alkaline Phosphatase > x2.5 normal
2. baseline renal impairment - serum creatinine > 200 µmol/l, eGFR < 30
3. cytopaenias - white cell count < 4.0, haemoglobin < 10, platelet < 150
4. pregnancy, planned pregnancy or breast feeding
5. contraindication to MRI
6. any co-morbid condition that would preclude the use of combination DMARD therapy

If you meet any patient with a new diagnosis of early rheumatoid arthritis who wishes to be considered for inclusion please contact Dr James Dale on either **0141 211 3008 (53008)** or **07855 823 209** or by emailing [james.dale@ggc.scot.nhs.uk](mailto:james.dale@ggc.scot.nhs.uk)

# Appendix C – Patient Information Leaflet

## Targeting Synovitis in Early Rheumatoid Arthritis (TASER)

We would like to invite you to participate in the **Targeting Synovitis in Early Rheumatoid Arthritis** study.

### **BACKGROUND**

Rheumatoid arthritis causes chronic inflammation in the joints, which results in pain, stiffness and restriction of movement. Over a period of time, this can lead to joint damage and restriction of day to day activities. Usually, patients with rheumatoid arthritis commence disease modifying drugs (methotrexate, sulfasalazine and/or hydroxychloroquine) to reduce any pain or stiffness and protect the joints from joint damage by dampening down the inflammation. Unfortunately, they do not cure the arthritis and many patients still experience some joint symptoms.

### **AIM**

We are trying to improve our treatment so that as many patients as possible experience complete control of their arthritis. If a single disease modifying drug does not work, adding one or more additional drugs can be effective (although this does mean taking more medicines). In this research trial we are exploring whether using an ultrasound scanner (similar to the type used to look at babies in the womb) will improve the way that we make decisions on the treatment of the early stages of rheumatoid arthritis.

### **YOUR PARTICIPATION**

We invite you to participate in this study and, if you agree, your treatment will be allocated randomly (by chance, like the toss of a coin) to either a standard assessment group or an ultrasound assessment group. You will be asked to attend clinic every month for 18 months. At each clinic visit we will take a further blood test (about two teaspoonfuls) and we will assess how active your arthritis is. If you are allocated to the routine group the activity of your arthritis will be measured by examining your joints. If you are allocated to the ultrasound group, you will still have your joints examined, but you may also have an ultrasound scan to look at the lining of the joint. This is a painless technique and usually takes about 15 minutes. In either group, if your arthritis is still active we will offer you additional treatment – either by increasing the dose of your tablets, by adding in a new treatment and/or by offering a steroid injection.

### **TREATMENT**

Initial treatment will follow guidelines that are similar to current practice using medications that are very commonly used to treat rheumatoid arthritis. You will initially be prescribed either methotrexate or sulfasalazine and we will discuss which you feel is the most suitable. If you continue to show evidence of joint inflammation we will prescribe a combination of methotrexate, sulfasalazine and hydroxychloroquine. At the end of the study you will likely continue on your final combination of methotrexate, sulfasalazine and/or hydroxychloroquine and your on going care will be transferred to the usual rheumatology follow-up clinic.

During the study period, if you continue to have joint inflammation despite the use of the methotrexate, sulfasalazine and hydroxychloroquine you will be offered treatment with etanercept, which is a 'biologic' or 'anti-TNF' drug. These drugs have been found to be more effective than ordinary disease modifying drugs. At present, the use of these drugs is restricted to patients who have very active disease, and they have usually been used in patients after several years of disease. However, there is evidence that they work well in patients with milder disease, and earlier on in the course of arthritis.

Patients in this trial could be offered etanercept at an earlier stage, and in milder disease than would otherwise be possible. If you are considered for etanercept you will need to undergo an additional chest x-ray and blood test before the treatment can commence. This is to find out if you

have ever been exposed to tuberculosis which, occasionally, can be reactivated by etanercept. Furthermore, as a result of its action on the immune system etanercept can make you more prone to other common infections. Therefore, if you are started on etanercept and develop a sore throat, a fever or other features of infection you should consult the investigators urgently. Patients who are doing very well on etanercept, and have no joint inflammation, will have their treatment discontinued after six months to see if the initial benefit persists. If the joint inflammation recurs you may be reconsidered for either etanercept or a similar medication.

You should remember that all drugs can cause side effects as well as benefits. Before you are prescribed any new disease modifying drug treatment, you will be given verbal and written information about the possible side effects and allowed time to discuss any questions that may arise. You are free to decline any drug treatment offered to you during the study, you are also free to withdraw from the study at any point. If you are prescribed methotrexate or sulfasalazine you will need to undergo regular blood tests (usually fortnightly at first) to monitor the effects of treatment. These blood tests are routine and are usually arranged through your General Practitioner's surgery

## **ASSESSMENTS**

At the start and the end of the trial we will take X-rays of your hands and feet and MRI scans of your dominant hand to determine what impact your treatment is having upon your arthritis' progress. In order to perform the MRI scans we will require you to attend for two additional appointments; one at the start and one at the end of the study. Every 3 months (for 18 months) you will be reviewed by a research nurse who will assess how active your arthritis is but will not know which group you have been allocated. You should not discuss with the nurse which treatment group you have been allocated to

## **ADDITIONAL TESTS**

In each group we will ask you to provide some extra blood tests (about 8 teaspoonfuls) at the very start, and at intervals throughout the study. Altogether, we could ask for extra blood samples up to 6 times throughout the course of the study. These additional blood tests will be performed during your regular clinic consultation. Furthermore, if we offer you a joint injection, and we drain some fluid from the joint, we will send the fluid for similar testing. All additional blood tests and fluid samples will be transported to the Translational Medicine Research Collaboration Core Laboratory in Dundee for analysis and storage. We plan to assess the levels of particular proteins and genetic factors to see if these help us predict which patients do best in the long term. All blood and fluid samples will be stored at the Translational Medicine Research Collaboration and could be used in future research projects, including genetic analysis. All stored samples will be anonymised and any new research projects will seek all relevant approvals and authorisations before starting

Your participation in this trial is entirely voluntary and may not be of direct benefit to you. The study could result in improved treatment for patients with similar problems. If you do not wish to participate in the study, or choose to withdraw at any point, the standard of your care will in no way be affected and you need not provide a reason. If you agree to take part in this study, your GP will be kept informed about your participation and all your treatments. The usual NHS complaints and compensation scheme will be available to you. You must not take part in this study if you are breast feeding, pregnant, planning to become pregnant or are not using a reliable method of contraception.

If you have any problems or queries please contact Dr Duncan Porter (0141 211 3262) or Dr James Dale (0141 211 3008 / 07855 823 209)

Kind regards

# Appendix D – Patient Consent Form

## Consent Form

Study Title: **Targeting Synovitis in Early Rheumatoid Arthritis – TASER**

**Please  
each stat**

1. I confirm that I have read and understood the information sheet (v6) for the above study and have had sufficient opportunity to ask questions .....  
.....
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason, and without my medical care or legal rights being affected .....  
.....
3. I understand that sections of my medical notes may be looked at by responsible individuals who are either working on this study or from regulatory authorities where it is related to my taking part in this study. I give permission for these individuals to have access to my medical records .....  
.....
4. I understand that my General Practitioner will be informed that I have agreed to participate in this study .....  
.....
5. I **do / do not** agree to the storage of any tissue samples obtained during the course of this study for future medical research purposes. Any future research will receive additional Research Ethics approval before commencing .....  
.....
6. I agree to take part in the above study .....  
.....

\_\_\_\_\_  
**Signature of Participant**

\_\_\_\_\_  
**Print Name**

\_\_\_\_\_  
**Date**

\_\_\_\_\_  
**Signature of Person  
taking Consent**

\_\_\_\_\_  
**Print Name**

\_\_\_\_\_  
**Date**

\_\_\_\_\_  
**Signature of Researcher**

\_\_\_\_\_  
**Print Name**

\_\_\_\_\_  
**Date**

## Appendix E – Recording Instrument for MSUS Findings

Disease Activity Score: \_\_\_\_\_

ESR \_\_\_\_\_; Tender JC \_\_\_\_\_; Swollen JC \_\_\_\_\_; Global VAS \_\_\_\_\_

		Right		Left	
		Synovial Hypertrophy	Doppler Signal	Synovial Hypertrophy	Doppler Signal
Hands	Index MCPj	(0-3)	(0-3)	(0-3)	(0-3)
	Index PiPj	(0-3)	(0-3)	(0-3)	(0-3)
	Middle MCPj	(0-3)	(0-3)	(0-3)	(0-3)
	Middle PiPj	(0-3)	(0-3)	(0-3)	(0-3)
	Radiocarpal	(0-3)	(0-3)	(0-3)	(0-3)
Feet	2 <sup>nd</sup> MTPj	(0-3)	(0-3)	(0-3)	(0-3)
	5 <sup>th</sup> MTPj	(0-3)	(0-3)	(0-3)	(0-3)
<b>Total Score</b>		(0-21)	(0-21)	(0-21)	(0-21)
<b>Total Index</b>		(0-7)	(0-7)	(0-7)	(0-7)

Active RA = total Power Doppler index  $\geq$  2 across all joints

**Grading:** per Szkudlarek et al. *Arthritis Rheum* 2003

*Synovial Hypertrophy* – non-compressible hypoechoic intra-articular area (synovial thickening)

- Grade 0 no synovial thickening
- Grade 1 minimal synovial thickening; filling the angle between the periarticular bones, without bulging over the line linking tops of the bones
- Grade 2 synovial thickening bulging over the line linking tops of the periarticular bones but without extension along the bone diaphysis
- Grade 3 synovial thickening bulging over the line linking tops of the periarticular bones and with extension to at least one of the bone diaphysis

*Doppler Signal* – extend of Power Doppler signal identified within the synovium

- Grade 0 no flow in the synovium
- Grade 1 single vessel signals
- Grade 2 confluent vessel signals in less than half the area of the synovium
- Grade 3 vessel signals in more than half the area of the synovium

# Appendix F: The Health Assessment Questionnaire

We are interested in learning how your illness affects your ability to function in daily life.

Please mark with an X the response which best describes your usual abilities OVER THE PAST WEEK:

**1. DRESSING AND GROOMING**

	Without ANY UNABLE Difficulty	With SOME Difficulty	With MUCH Difficulty	to do
--	-------------------------------	----------------------	----------------------	-------

Are you able to:

a Dress yourself, including shoelaces and buttons? 1  2  3  4

b Shampoo your hair? 1  2  3  4

**2. ARISING**

	Without ANY UNABLE Difficulty	With SOME Difficulty	With MUCH Difficulty	to do
--	-------------------------------	----------------------	----------------------	-------

Are you able to:

a Stand up from an armless straight chair? 1  2  3  4

b Get in and out of bed? 1  2  3  4

**3. EATING**

	Without ANY UNABLE Difficulty	With SOME Difficulty	With MUCH Difficulty	to do
--	-------------------------------	----------------------	----------------------	-------

Are you able to:

a Can you cut meat? 1  2  3  4

b Lift a full cup or glass to your mouth? 1  2  3  4

c Open a new carton of milk? 1  2  3  4

**WALKING**

	Without ANY UNABLE Difficulty	With SOME Difficulty	With MUCH Difficulty	to do
--	-------------------------------	----------------------	----------------------	-------

Are you able to:

a Walk outdoors on flat ground? 1  2  3  4

b Climb up five steps? 1  2  3  4

**4. Please mark with an X any AIDS OR DEVICES that you usually use for any of these activities:**

1  Walking stick 5  Devices used for dressing (button hook, zipper pull, long handled shoe horn, etc.)

2  Walker 6  Built-up or special utensils

3  Crutches 7  Special or built-up chair

4  Wheelchair

8  Other

**5. Please mark with an X any categories for which you usually need HELP FROM ANOTHER PERSON:**

9  Dressing and Grooming

11  Eating

10  Arising

12  Walking

Please mark with an X the response which best describes your usual abilities OVER THE PAST WEEK:

**6. HYGIENE**

**Without  
ANY  
Difficulty**

**With  
SOME  
Difficulty**

**With  
MUCH  
Difficulty**

**UNABLE  
to do**

Are you able to:

a Wash and dry your body? 1

2

3

4

b Take a tub bath? 1

2

3

4

c Get on and off the toilet? 1

2

3

4

**7. REACH**

**Without  
ANY  
Difficulty**

**With  
SOME  
Difficulty**

**With  
MUCH  
Difficulty**

**UNABLE  
to do**

Are you able to:

a Reach out and get down a 5-pound object (such as a bag of potatoes) from just above your head? 1

2

3

4

b Bend down to pick up clothing from the floor 1

2

3

4

**8. GRIP**

**Without  
ANY  
Difficulty**

**With  
SOME  
Difficulty**

**With  
MUCH  
Difficulty**

**UNABLE  
to do**

Are you able to:

a Open car doors? 1

2

3

4

b Open jars which have been previously opened? 1

2

3

4

c Turn taps on and off? 1

2

3

4

**9. ACTIVITIES**

**Without  
ANY  
Difficulty**

**With  
SOME  
Difficulty**

**With  
MUCH  
Difficulty**

**UNABLE  
to do**

Are you able to:

a Run errands and shop? 1

2

3

4

b Get in and out of a car? 1

2

3

4

c Do chores such as vacuuming or garden work? 1

2

3

4

**10. Please mark with an X any AIDS OR DEVICES that you usually use for any of these activities:**

- |                            |   |                            |                                     |
|----------------------------|---|----------------------------|-------------------------------------|
| 1 <input type="checkbox"/> | Raised toilet seat                      | 4 <input type="checkbox"/> | Bathtub bar                         |
| 2 <input type="checkbox"/> | Bathtub seat                            | 5 <input type="checkbox"/> | Long-handled appliances for reach   |
| 3 <input type="checkbox"/> | Jar opener (for jars previously opened) | 6 <input type="checkbox"/> | Long-handled appliances in bathroom |
|                            |   | 7 <input type="checkbox"/> | Other                               |

**11. Please mark with an X any categories for which you usually need HELP FROM ANOTHER PERSON:**

- |                            |         |                            |                             |
|----------------------------|---------|----------------------------|-----------------------------|
| 1 <input type="checkbox"/> | Hygiene | 3 <input type="checkbox"/> | Gripping and opening things |
| 2 <input type="checkbox"/> | Reach   | 4 <input type="checkbox"/> | Errands and chores          |

## Appendix G: The EQ-5D Questionnaire

By placing a tick in one box in each group below, please indicate which statements best describe your own health state today.

### Mobility

- I have no problems in walking about
- I have some problems in walking about
- I am confined to bed

### Self-Care

- I have no problems with self-care
- I have some problems washing or dressing myself
- I am unable to wash or dress myself

### Usual Activities (*e.g. work, study, housework, family or leisure activities*)

- I have no problems with performing my usual activities
- I have some problems with performing my usual activities
- I am unable to perform my usual activities

### Pain/Discomfort

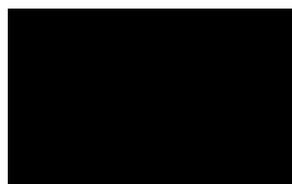
- I have no pain or discomfort
- I have moderate pain or discomfort
- I have extreme pain or discomfort

### Anxiety/Depression

- I am not anxious or depressed
- I am moderately anxious or depressed
- I am extremely anxious or depressed

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.



## Appendix H: Materials for Transcriptomic Analysis

The following lists describe the materials required to process and analyse a single PAXgene RNA sample using the methods described in sections 2.7.2.1 – 2.7.2.3

### *Standard Laboratory Equipment*

- Centrifuge
- Single chamber pipette: volumes ranging from 1µl to 4ml
- Incubation unit and thermal cycler
- Orbital shaker
- Rocker mixer
- Hybex waterbath
- Staining dish
- Slide rack holder
- Fine forceps and tweezers
- Nanodrop 1000 spectrophotometer
- Agilent Technologies Bioanalyser

### *RNA Concentration and Purification*

- PAXgene RNA tube containing whole blood
- PAXgene spin column
- PAXgene shredder spin column
- PAXgene processing tubes
- RNase-free water – 4ml
- Proteinase K - 350µl
- Re-suspension buffer - 350µl
- Binding buffer - 300µl
- 96-100% Ethanol - 350µl
- Wash buffer
- Elution buffer - 40µl
- DNase incubation mixture:           10µl DNase I  
  70µl DNA digestion buffer

### *cDNA Hybridisation and RNA Amplification*

- Nuclease-free water
- Reverse transcription master mix: 1µl T7 Oligo(dT) primer  
  2µl 10X first strand buffer  
  4µl dNTP mix  
  1µl RNase inhibitor  
  1µl Arrayscript
- Second strand master mix:           63µl nuclease-free water  
  10µl 10X second strand buffer  
  4µl dNTP mix  
  2µl DNA polymerase  
  1µl RNase HI
- cDNA binding buffer - 250µl
- cDNA filter cartridge
- Wash buffer
- IVT master mix:                        2.5µl T7 10X reaction buffer  
  2.5µl T7 enzyme mix  
  2.5µl Biotin-NTP mix
- cRNA binding buffer - 350µl
- ACS reagent grade 100% ethanol - 250µl
- cRNA filter cartridge

### *Illumina Beadchip Microarray Analysis*

Illumina HumanHT-12v4.0 Beadchip microarray chip

Illumina BeadArray Reader

Hybridisation chamber: comprising hybridization chamber, chamber gasket, chamber insert and lid

Hybridisation tube

Beadchip wash tray and cover

Hybridisation buffer - 10 $\mu$ l

Humidity control buffer – 200 $\mu$ l

RNase-free water

E1BC buffer – 3ml (dissolved in 1L RNase-free water)

100% ethanol – 250ml

Block E1 buffer – 6ml

Cy3-Streptavidin – 1: 1,000 dilution

