The impact of achieving low disease activity in the first year of disease on future disability and damage in early rheumatoid arthritis

by

Pooneh Seyed Akhavan, MD

A thesis submitted in conformity with the requirements
for the degree of Masters of Science

Health Policy, Management and Evaluation
University of Toronto

The impact of achieving low disease activity in first year of disease on disability and damage in patients with early rheumatoid arthritis

Pooneh Seyed Akhavan Masters of Science (Clinical Epidemiology)

Institute of Health Policy, Management and Evaluation,

University of Toronto, 2013

Abstract

Aim: To describe the predictive validity of reaching low disease activity (LDA) at 1 year on future disability and joint damage in patients with early rheumatoid arthritis (ERA).

Methods: First a systematic literature review of prognostic studies assessing the association between disease activity and functional or radiographic outcomes in ERA was performed. Then data from the Study Of New-Onset RA (SONORA) were used to evaluate the impact of year-one LDA on 3-year disability and 2-year radiographic progression using multivariate regression analyses.

Results: Our review demonstrated evidence for relationship between baseline disease activity and future disability and join damage. However evidence for the impact of early treatment response on long-term outcomes in ERA is sparse. Analysis of 984 patients showed year one LDA predicts lower HAQ ($p<.0001$) and less damage ($p=0.04$) in future.

Conclusion: Reaching LDA early is associated with better long-term functional and radiographic outcomes in patients with early RA.

Acknowledgements

I would like to thank my thesis supervisor, Dr. Claire Bombardier for her guidance, kind support and encouragement throughout my graduate training and her constant invaluable inspiration for my professional development. I would like to extend my appreciation to my thesis committee members Dr. Paul Fortin and Dr. George Tomlinson for their helpful comments, criticisms, and broad expertise. I greatly appreciate their enthusiasm, guidance and valuable suggestions.

I am also grateful to the research staff at Toronto General Hospital including Dr. Maggie Chen and Dr. Xining Li.who helped me at different stages of my project. I would like to express my gratitude to the Arthritis Society, Canadian Rheumatology Association and UCB Canada for supporting me with their fellowship award and to the Arthritis Research Foundation for providing additional financial support during my training. I would like to thank all my colleagues at Rebecca McDonald Centre for Arthritis and Autoimmune disease at Mount Sinai Hospital for their support and encouragement.

I thank my loving and supportive family, my parents and my sisters for their encouragement. Finally, I dedicate this thesis to, my husband, Farshid. Your endless love, sense of humor, unwavering support for my career, and constant encouragement are the only reason I can do what I do. Thank you.

Table of Content **Page**

1 Introduction and thesis overview 1
2 The association between disease activity and long-term outcomes in patients with early rheumatoid arthritis - a systematic literature review
2.1 Introduction 3
2.2 Methods 4
2.2.1 Literature search 4
2.2.2 Study selection 4
2.2.3 Data abstraction and synthesis 6
2.3 Results 6
2.3.1 Functional outcome 6
2.3.2 Radiographic outcome 12
2.4 Discussion 15
3 The impact of reaching low disease activity in the first year of disease on disability or joint damage in patients with early rheumatoid arthritis- results from a multicenter cohort
3.1 Introduction 22
3.2 Study objectives 23
3.3 Methods 23
3.3.1 Study design 24
3.3.2 Study population 24
3.3.3 Study plan 25
3.3.4 Research ethics approval 26
3.3.5 Data collection 27
3.3.6 Main outcomes (dependent variables) 27
3.3.7 Main predictor (independent variable) 30
3.3.8 Sample size 31
3.4 Analysis 31
3.4.1 Statistical software 31
3.4.2 Descriptive statistics 31
3.4.3 Model building 32
3.4.4 Missing data 34
3.5 Results
3.5.1 Study population 40
3.5.2 Baseline characteristics 40
3.5.3 Description of main predictor and outcome 42
3.5.4 Achieving LDA at year 1 and functional outcome 44
3.5.5 Achieving LDA at year 1 and radiographic damage 48

Table of Content **Page**

3.5.6 Exploratory analysis 52
3.6 Discussion 54
3.6.1 Key findings and clinical implications 54
3.6.2 Study weaknesses 58
3.6.3 Study strength 59
3.6.4 Unanswered questions and future research 60
4 Conclusions 63
5 References 64
6 Appendices 75

List of Tables — Page

Table 2-1: Prognostic studies. Outcome = Function 8

Table 2-2: Final predictive models - Outcome: Functional limitation (HAQ) 9

Table 2-3: Prognostic studies. Outcome = Radiographic joint damage 14

Table 3-1: Baseline characteristics 41

Table 3-2: Disease activity (SDAI), HAQ and Sharp Score at different time points 42

Table 3-3: Disability severity at baseline 43

Table 3-4: Comparison of functional and radiographic outcomes in two groups 44

Table 3-5: Univariate regression analysis. Outcome = HAQ year 3 (n=984) 45

Table 3-6: Test for multicollinearity (Primary outcome: Functional disability) 46

Table 3-7: Linear regression models. Complete case analysis after transformation 47

Table 3-8: Linear regression models. Analysis after multiple imputations (n=984) outcome: HAQ at year 3 48

Table 3-9: Univariate regression analysis. Outcome: X-ray progression at 2 years 49

Table 3-10: Logistic regression models. Outcome: 2 year x-ray progression. Complete case analysis (n=485) 50

Table 3-11: Logistic regression models. Outcome: x-ray progression at year 2. Imputed data analysis 51

List of figures Page

Figure 2-1: Results of the literature search and disposition of the potentially relevant studies 7

Figure 3-1: The multiple imputation process using SAS software 37

Figure 3-2: HAQ pattern over 3 year follow up 43

List of appendices **Page**

Appendix 1-A 75
Appendix 1-B 78
Appendix 1-C 81
Appendix 1-D 84
Appendix 2-A 94
Appendix 2-B 95
Appendix 2-C 96
Appendix 2-D 97
Appendix 2-E 100
Appendix 2-F 101
Appendix 2-G 102
Appendix 2-H 104
Appendix 2-I 106
Appendix 2-J 108
Appendix 2-K 110
Appendix 2-L 112

Glossary of acronyms

ACR	American college of rheumatology
Anti-CCP	anti-cyclic citrullinated peptide
CS	consensus statement
CPG	clinical practice guideline
CRP	c-reactive protein
DAS28	disease activity score based on 28 joint count
ESR	erythrocyte sedimentation rate
ERA	early rheumatoid arthritis
EULAR	European league against rheumatism
HAQ	health assessment questionnaire
LDA	low disease activity
MI	multiple imputation
MTX	methotrexate
RA	rheumatoid arthritis
RF	rheumatoid factor
SDAI	simplified disease activity index
SJC	swollen joint count
TJC	tender joint count
TJSN	total joint space narrowing score
TERO	total erosion score

1 Introduction and thesis overview

The treatment of rheumatoid arthritis (RA) has changed tremendously over the past two decades. Many clinical practice guidelines (CPG) and consensus statements (CS) based on systematic reviews and the collective opinion of experts have been developed to assist rheumatologists in clinical decision making and to improve quality of care and eventually patient outcomes [1-5].

Recent guidelines recommend that the primary target for treatment of RA should be a state of clinical remission but low disease activity (LDA) may be an acceptable alternative therapeutic goal, particularly in established long-standing disease [2-4]. This recommendation is based on strategic trials demonstrating better outcomes when treatment was aimed to attain a LDA state [4, 6] and trials in which tight control was mandatory in all treatment arms, but the protocolized therapies differed [6, 7]. These trials mainly showed superior disease activity outcome (e.g. number of patients in remission) in the intensive treatment group while functional or radiographic outcomes were not assessed in many of these trials and when assessed were not significantly different in all [6].

The strategic trials were focused on comparing routine care to an intensive protocol which is not usually feasible to incorporate in daily practice (e.g. monthly assessment and intra-articular injection of any swollen joint with corticosteroids at each visit and treatment escalation when target was not reached) and provided supporting evidence for improved outcomes at the end of the follow-up at the group level in patients who were treated aggressively. Assessment of long-term outcomes was not the objective of these studies and the impact of attaining the target on

future outcomes was not addressed. It still remains to be answered whether achieving the desired treatment response, LDA, is associated with improved long-term outcomes in patients with early disease.

Assessment of associations between risk factors/predictors and health outcomes can be accomplished by analyzing longitudinal data using prognostic modeling. The aim of the current study was to evaluate the predictive validity of early response, achieving LDA, on long-term outcomes in patients with early RA. We chose to analyze data from an observational cohort of early RA patients which was established before the concept of treat to target was introduced. Hence, patients were treated based on usual care and there should not be any significant differences in treatment strategies among participating rheumatologists.

This thesis includes four chapters. Chapter 2 lays out a systematic literature review which was conducted first to collect published evidence assessing the association between disease activity and two main long-term outcomes in RA, disability and radiographic damage, in prognostic studies. It includes the background, methods, results and a discussion of the findings. In chapter 3 the analysis of data from the Study Of New Onset Rheumatoid Arthritis (SONORA) is presented. SONORA is a cohort of patients with early rheumatoid arthritis and a prognostic analysis was performed to assess the impact of reaching LDA on disability and joint damage in this patient population. This chapter includes the background, study objectives, methods, results, a discussion of findings, limitations and strengths of the analysis and finally a summary and conclusion is presented in chapter 4.

2 The association between disease activity and long-term outcomes in patients with early RA - a systematic literature review

2.1 Introduction

Rheumatoid Arthritis (RA) is an auto-immune disease characterized by chronic inflammation which can result in destruction of the joints and disability [8]. Active inflammatory disease is considered the process that leads to joint damage and functional limitation and therefore is the main target of RA treatment [9]. Given the potential consequences of the inflammatory changes of RA, interference with the active disease process is essential. A number of investigator-initiated trials have shown that rapid switching of therapy upon missing a targeted disease activity state, remission or LDA, will lead to superior outcomes when compared with routine, unsystematic monitoring and change of therapy [6, 10, 11].

The ultimate goal is to prevent future radiological joint damage and functional disability in patients with RA. However, the relationship between disease activity and long-term outcomes in early disease remains a topic of debate. Some studies have suggested that radiographic damage may progress independently of disease activity [12, 13] and some suggested a significant association [14-18] .

In 2010, Van Tuyl et al performed a systematic review of the literature to assess the relationship between remission and long-term outcomes in patients with rheumatoid arthritis and showed an association [19]. This review was limited to the studies that reported remission and relied on a wide spectrum of study designs.

To our knowledge there is no systematic literature search reviewing published evidence on the longitudinal association between disease activity and long-term outcomes in RA. Patients with early disease are the main target of early, aggressive treatment. The goal of our review was to collect existing prognostic studies assessing the impact of treatment response and disease activity on future joint damage and disability in patients with early RA.

2.2 Methods

2.2.1 Literature search

A systematic search of the published literature was conducted using MEDLINE (1945 to February week 5 2012) and EMBASE (1980 to 2012 week 09) databases. The search was limited to studies published in English focused on four components: RA, low disease activity, joint damage and disability (Appendix 1-A). Reference lists of selected relevant studies from the electronic search were manually searched to identify additional eligible studies.

2.2.2 Study selection

First, titles and abstracts of all retrieved references were reviewed, excluding articles that were clearly not pertinent. Based on the initial protocol only papers evaluating the impact of low disease activity on disease outcomes were to be included however due to small number of eligible papers and after further discussion with the thesis committee it was decided to include all papers that had evaluated the prognostic impact of disease activity (all levels) and above mentioned outcomes. Second, the full text of selected articles was reviewed.

Inclusion criteria:

1) Included adult (age ≥ 18 years) RA patients with disease duration ≤24 months;

2) Reported "joint damage: as the outcome, measured on x-ray with Larsen, Sharp or van der Heijde/Sharp score OR reported "disability" as the outcome, measured with one or more of the following: HAQ, arthritis impact measurement scale, RA quality of life;

3) Assessed the disease activity as a predictor using a valid measure such as composite indices (e.g. DAS28) or active joint count (e.g. swollen joint count)

4) Longitudinal studies (including RCTs);

5) Statistical analysis: studies had to use multivariate analysis, e.g., multiple logistic regression or Cox proportional hazards model, to identify the potential predictive impact of disease activity on long-term outcome while adjusting for potential confounders.

Exclusion criteria:

1) Association between the disease activity measure and damage/disability was estimated cross-sectionally;

2) Reviews, case reports, case series, editorial and comments/letters;

2.2.3 Data abstraction and synthesis

Data abstraction forms were developed to summarise the studies. We summarized the data by stratifying outcomes into 2 groups:

1) Physical function and
2) Joint damage

We further categorized the second group into 3 subgroups based on the outcome:

a) Outcome was joint damage score at end point (continuous variable)

b) Outcome was the change in damage score (Δscore) from baseline to end point (continuous variable)

c) Outcome was x-ray progression (significant Δscore) at end point (dichotomous variable).

The association between predictors and outcome was abstracted as presented in the study [i.e. odds ratio (OR), hazard ratio (HR), or coefficient (β), with its corresponding 95% confidence interval (95% CI)]. Data on study characteristics and prognostic model was extracted and presented in tables.

2.3 Results

After reviewing titles and abstracts of 2285 citations, 117 papers were selected for full review. Thirty-seven studies were identified using the prespecified inclusion and exclusion criteria. Bibliographies of selected articles were scanned and three additional papers were included after this review (Figure 2-1).

Forty articles included in this review, eight addressed predictors of functional outcome [20-27] , 31 addressed radiographic damage [14-18, 28-53] and one assessed both outcomes [54] (Appendix 1-B).

2.3.1 Functional outcome

Among studies assessing functional outcome, all except 3 showed a statistically significant association between disease activity and function (see Tables 2-1 and 2-2). Follow up duration varied from 3 months to 10 years. Baseline/prior HAQ was a significant predictor of HAQ at end point in most studies. Older age and female sex were other significant predictors of disability [20, 21, 54]. Joint damage score remained significant in 2 studies [20, 21].

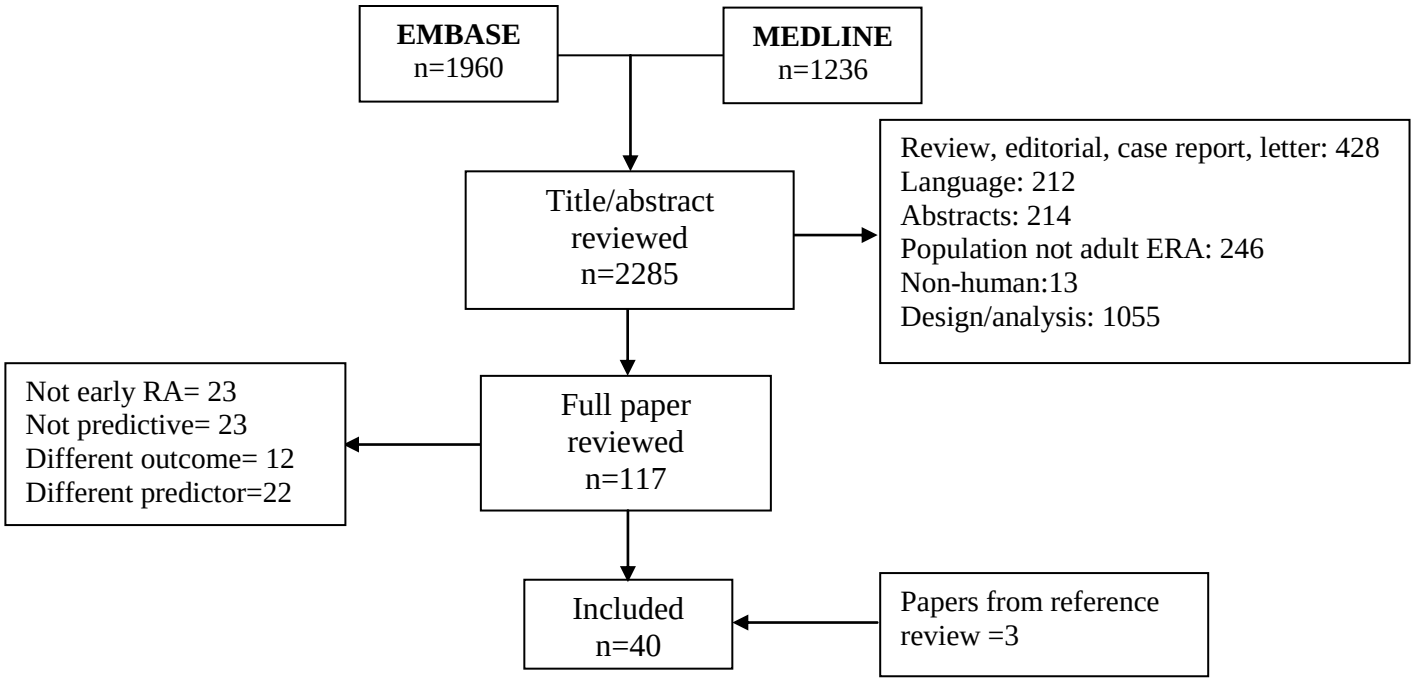

Figure 2-1-Results of the literature search and disposition of the potentially relevant studies

Two studies analyzed data from the BeST trial [24, 26]. Dirven et al. assessed predictors of short-term (3 months) HAQ in this early RA population using logistic regression analysis [24]. They demonstrated that baseline HAQ, pain, Ritchie Articular Index (RAI) and treatment group were significant independent predictors of poor functional outcome (HAQ>1) at 3 months whereas previously known risk factors of joint damage including baseline damage, RF and anti-CCP were not. Van der Kooi et al performed a longitudinal data analysis on 5 year follow-up data from the BeST study to evaluate the correlation between DAS and HAQ over 5 years. In this cohort, HAQ had improved during the 5 year follow up. They showed that a decrease in

DAS is associated with a decrease in HAQ which depends on the ΔDAS and the absolute DAS level but is independent of follow-up duration [26].

Table 2-1: Prognostic studies. Outcome = Function

Study	Year	Country	Design (Cohort/RCT)	Sample size	Disease duration	F/u duration	Recruitment years
1 Dirven	2012	Netherlands	RCT (BeST)	497	<2yr	3 mo	2000-02
2 van der Kooi	2011	Netherlands	RCT (BeST)	508	<2yr	5 yr	2000-02
3 Verstappen	2007	Netherlands	RCT*	112	<1yr	7 yr	1999-2003**
4 Bansback	2006	UK	Cohort (ERAS)	985	<2 yr	5 yr	1986
5 Combe	2003	France	Cohort	191	<1yr	5 yr	1993-94
6 Lindqvist	2002	Sweden	Cohort	183	<2yr	10 yr	1985-89
7 Welsing	2001	Netherlands	Cohort	203	<1 yr	6 yr	1985-98
8 Kroot	2000	Netherlands	Cohort	237	<1yr	6 yr	1985
9 Corbett	1993	UK	Cohort	102	<1yr	5 yr	1966-71

*Long-term follow-upstudy of 2 inception cohorts from 2 trials: CAMERA and ERA, ** for CAMERA study, F/U=follow up, ERAS=Early RA Study,

Verstappen et al analyzed data from 2 inception cohorts of patients with early RA (follow-up cohorts of two 2-year trials) and showed the only significant baseline predictor of HAQ limitation at 7 years was worse functional disability (OR, 95% CI: 2.5, 1.30-5.32) [27]. Thompson score (active joint score) at baseline was not significant in this analysis.

Table 2-2: Final predictive models- Outcome: Functional limitation (HAQ)

No	Study	Year	Analysis	Outcome	predictors	Estimate(95%CI)	OR (95%CI)	p-value
1	**Dirven**	**2012**	Logistic regression	Short-term disability	Rx: Monotherapy		Reference	
					Comb w Prednisone		0.3 (0.2-0.5)	
					Comb w Infliximab		0.4 (0.2-0.6)	
					HAQ : <1.38		Reference	
					1.38-2		2.6 (1.6-4.2)	
					>2		5.3 (2.9-9.5)	
					VAS pain: <40		Reference	
					40-60		2.2 (1.3-3.8)	
					>60		2.7 (1.4-5.1)	
					RAI: <10		Reference	
					10-16		1.7 (1.0-2.9)	
					>16		2.7 (1.5-4.7)	
2	**Van der Kooi**	**2011**	LMM	HAQ over 5 years	Ln time	0.044 (0.031-0.057)		
					Previous HAQ	0.234 (0.213- 0.255)		
					Previous DAS	0.213 (0.200-0.226)		
					Delta DAS	0.183 (0.166- 0.200)		
					PreviousDAS x ΔDAS	0.022 (0.016-0.027)		
3	**Verstappen**	**2007**	Logistic regression	HAQ limitation, 7 y	Worst functional disability BSL		2.63(1.30-5.32)	0.007
4	**Bansback**	**2006**	Logistic regression	mod-severe *vs No/mild disability at 5 yr	DAS28 y 1	0.138	1.148	0.064
					HAQ BSL	0.532	1.702	0.005
					HAQ y 1	0.894	2.445	0.000
					Larsen BSL	0.013	1.013	0.200

No	Study	Year	Analysis	Outcome	predictors	Estimate(95%CI)	OR (95%CI)	p-value
					Hb BSL	0.015	1.015	0.025
5	**Combe**	**2003**	Linear regression	HAQ at yr 5	HAQ BSL	0.394		0.0001
					ESR BSL	0.008		0.006
					CRP BSL	0.005		0.001
					RAI BSL	0.021		0.045
6	**Lindqvist**	**2002**	Logistic regression	HAQ>1 , 10 yr	Mean HAQ in 3mo		13.36(5.08- 35.14)	
7	**Welsing**	**2001**	GLMM	HAQ over 6 yrs	Higher SHS	0.005(0.002-0.008)		
					Higher DAS	0.10 (0.07-0.13)		
					Higher age	0.01 (0.01-0.20)		
					RF +	0.19 (0.03-0.35)		
					Female sex	0.22 (0.08-0.36)		
					Transformed terms			
					Higher sSharp	-0.08 (-0.12, -0.03)		
8	**Kroot**	**2000**	Linear regression	HAQ at 3yrs	Female	-0.168(0.052)* £		
					Age (yrs)	0.006(0.002)***		
					RF +	0.185(0.065)*		
					DAS	0.105(0.022)**		
					HLA-DR4 +	0.013(0.051)		
					Anti-CCP +	-0.024(0.053)		
				HAQ at 6yrs	Female	-0.128(0.055)**		
					Age (yrs)	0.009(0.002)***		
					RF +	0.150(0.070)**		

No	Study	Year	Analysis	Outcome	predictors	Estimate(95%CI)	OR (95%CI)	p-value
					DAS	0.100(0.024)***		
					HLA-DR4 +	<0.001(0.055)		
					Anti-CCP +	-0.002(0.056)		
				ΔHAQ at 3yrs	Female	-0.074 (0.059)		
					Age (yrs)	0.004(0.002)**		
					RF +	0.028(0.072)		
					DAS	-0.086(0.025)*		
					HLA-DR4 +	0.065(0.058)		
					Anti-CCP +	-0.033(0.063)		
				ΔHAQ at 6yrs	Female	-0.027(0.065)		
					Age (yrs)	0.008(0.002)***		
					RF +	<0.001(0.077)		
					DAS	-0.086(0.027)*		
					HLA-DR4 +	0.018(0.064)		
					Anti-CCP +	-0.015(0.068)		
9	**Corbett**	**1993**	Discrimin--ant analysis	HAQ at 5 years	AJC during first 2yr			p<0.05
				Functional group	Erosion in 2 years			p<0.05
				I , II or III/IV	Poor grip strength			p<0.05
					RF +			p<0.05
					Increased age			P<0.05
					High body mass			P<0.05

*p<0.005, **p<0.05, ***p<0.0001, £: β (standard error); comb w=combination with; LMM=linear mixed model; DAS=disease activity score, RF=rheumatoid factor; SHS=Sharp van der Heijdeh Score; RAI=Ritchie Articular Index; Hb=Hemoglobin; BSL=baseline;

The remaining studies used data from other early RA cohorts [20-23, 25, 27, 54, 55] with follow-up duration of 5 to 10 years. Study participant population varied from 102 [20] to 985 [22]. Except one [25], all showed a significant association between disease activity measures and function (Table 2-2). Lindqvist et al, evaluated data from a cohort of 183 early RA patients who were followed annually for 10 years. The mean HAQ had increased from 0.8 to 1.1 during the follow-up and the only significant predictor of HAQ at end point was the mean HAQ during the first 3 months (OR, 95% CI: 13.36, 5.08-35.14) [25]. In the analysis performed by Bansback et al on 985 patients enrolled in the Early Rheumatoid Arthritis Study (ERAS), they showed that higher DAS28 at year 1 was associated with worse functional outcome at 5 years but this association was not statistically significant (OR=1.148, p=0.064) [22]. The four other studies demonstrated a significant association between disease activity measures and functional outcome (Table 2-2).

2.3.2 Radiographic outcome

2.3.2.1 Outcome: Joint damage score (continuous variable)

The main outcome was joint damage score at end point in five studies [15, 44, 51, 52, 54]. Radiographic outcome was assessed at 2 years in all except in one study where it was evaluated at 3 and 6 years [54]. Linear regression was used for analysis. Three studies showed a significant association between disease activity measure at baseline and damage score at follow-up (Table 2-3, Appendices 1-C, 1-D).

2.3.2.2 Outcome: Change in joint damage score (continuous variable)

Change in joint damage score was evaluated as the main outcome in 14 studies [14, 16, 28, 31, 36-39, 41, 43, 46, 52-54] (Table 2-3, Appendices 1-C & 1-D). Sample size varied from 43 to 336 and follow-up duration was up to 9 years. Five studies did not show any significant association between baseline disease activity and x-ray score progression [28, 36, 37, 39, 41, 46]. Hetland et al assessed the association of radiographic damage and clinical factors in patients enrolled in CIMESTRA trial at the end of the 2-year trial and after 5 years. They showed that bone marrow edema on MRI was an independent predictor of joint damage at both time points and disease activity at baseline was not [36, 37]. Boyensen et al. also demonstrated that baseline bone marrow edema and synovitis on MRI were independent predictors of 3-year radiographic progression in an inception cohort of early RA patients but clinical measures of disease activity were not significant [28]. Nyhall-Wahllin et al adjusted their prognostic model assessing the impact of RA nodule on joint damage for potential confounders and in this analysis, baseline damage score and positive anti-CCP were the only significant factors [46]. This result is similar to Kaltenhauser et al paper that also performed multivariate linear regression analysis predicting joint damage progression at 2 years and showed a significant association between damage progression, baseline damage, RF and shared epitope but not the baseline and 6 month disease activity measures [39]. The remaining 9 studies showed that higher disease activity was associated with worse radiographic outcome and a good treatment response at 6 months predicted less progression [14, 16, 31, 38, 41, 43, 52-54] (Appendix 1-D)

Table 2-3: Prognostic studies. Outcome = Radiographic joint damage* studies that showed a significant association between disease activity and damage are highlighted (grey)

No	Study	Year	Country	Design (cohort, RCT)	Sample size	Disease duration	F/U duration
	Outcome: Joint damage score (continuous variable)						
1	Berglin	2006	Sweden	Cohort	138	<1yr	2 yr
2	Kroot	2000	Netherlands	Cohort	237	<1yr	6 yr
3	van der Heijdeh	1992	Netherlands	Cohort	147	<1yr	2 yr
4	Manfredsdottir	2006	Iceland	Cohort	100	<1yr	2 yr
5	Tengstrand	2004	Sweden	Cohort (BARFOT)	844	<1yr	2 yr
	Outcome: Joint damage score change (continuous variable)						
1	Bakker	2011	Netherlands	RCT (CAMERA)	299	<1 yr	5 yr
2	Ichikawa	2009	Japan	Cohort	55	<2yr	8 yr
3	De Vries-Bouwstra	2006	Netherlands	Cohort (Leiden)	152	<2yr	1 yr
4	Machold	2007	Austria	Cohort (VERA)	138	<3mo	3 yr
5	Welsing	2004	Netherlands	Cohort**	185+152	<1yr	9, 6 yr
6	Berglin	2003	Sweden	ERA cohort	43	<1yr	2 yr
7	Landewe	2002	Netherlands	RCT (COBRA)	115	<2yr	4-7 y
8	Kroot	2000	Netherlands	Cohort	237	<1yr	3 yr
9	van der Heijdeh	1992	Netherlands	Cohort	147	<1yr	2 yr
10	Nyhall-Wahlin	2011	Sweden	Cohort (BARFOT)	336	<1yr	5 yr
11	Boyensen	2011	Norway	Cohort	84	<1yr	3 yr
12	Hetland	2010	Denmark	RCT (CIMESTRA)	130	<6mo	5 yr
13	Hetland	2009	Denmark	RCT (CIMESTRA)	160	<6mo	2 yr
14	Kaltenhauser	2001	Germany	Cohort	87	<2yr	2 yr
	Outcome: Joint damage progression (dichotomous variable)						
1	Salaffi	2011	Italy	Cohort	59	<1yr	3 yr
2	Westhoff	2008	Germany	Cohort	896	<2yr	3 yr
3	Berglin	2006	Sweden	Cohort	138	<1yr	2 yr
4	Mottonen	1998	Finland	Cohort	142	<2yr	6 yr
5	Mouterde	2011	France	Cohort (ESPOIR)	736	<6mo	6 mo
6	Hetland	2010	Denmark	RCT (CIMESTRA)	130	<6mo	5 yr
7	Courvoisier	2008	France	Cohort	191	<1yr	10 yr
8	Sanmarti	2007	Spain	Cohort	105	<2yr	2 yr
9	Tanaka	2005	Japan	Cohort	130	<1yr	10 yr
10	Dixey	2004	UK	Cohort (ERAS)	866	<2yr	3 yr
11	Forslind	2004	Sweden	Cohort (BARFOT)	379	<1yr	2 yr
12	Goronzy	2004	USA	Cohort	111	<1yr	2 yr
13	Korpela	2004	Finland	RCT (FIN-RACo)	195	<2yr	5 yr
14	Lindqvist	2003	Sweden	Cohort	183	<2yr	10 yr
15	Sanmarti	2003	Spain	Cohort	60	<2yr	1 yr
16	Combe	2001	France	Cohort	191	<1yr	3 yr
17	Fex	1996	Sweden	Cohort	113	<2yr	5 yr

F/U=follow up, ERA=early rheumatoid arthritis, yr=year, mo=month(s), VERA=Very early RA cohort; **Included two groups from long-term extension of COBRA trial and University Medical Center Nijmegen (UMCN) cohort;

2.3.2.3 Outcome: x-ray progression (dichotomous variable)

Seventeen studies analyzed x-ray progression of joints as a dichotomous outcome using logistic regression analysis (Table 2-3). Patient population varied from 48 up to 896 and follow-up duration was as short as 6 months up to 10 years. There was a significant heterogeneity among outcome definitions in the studies included (Appendix 1-C).

Only 4 studies showed a significant association between the disease activity measures and radiographic progression [15, 17, 18, 47]. Salaffi et al analyzed data from a small (n=48) cohort of early RA patients and showed time-integrated DAS28-CRP was a significant predictor of change in Sharp van der Heijde Score (ΔSHS) > 9.5 (defined as the smallest detectable difference, SDD) [47]. Westhoff et al found that patients with high disease activity (DAS28>5.1) at baseline were more likely to progress radiologically at 3 years compare to patients in low disease activity status or in remission (DAS28 <3.2) (OR, 95% CI: 1.9 (1.1, 3.1)) [47]. Two studies demonstrated better radiographic outcome at 2 years in patients who had a moderate or good treatment response at 6 months compare to patients who had a poor response [16, 17].

The remaining studies did not find any significant association between baseline disease activity and radiographic outcome. Baseline joint damage and auto-antibodies (RF and anti-CCP) were significant independent predictors of radiographic progression in most studies (Appendix 1-D).

2.4 Discussion

We systematically reviewed prognostic studies that evaluated the impact of disease activity on two main outcomes, function and radiographic damage, in patients with early RA

(disease duration ≤ 2 years). We included longitudinal studies that used a multivariate analysis and identified 40 eligible studies.

Published evidence suggests that a correlation exists between disease activity and future functional capacity. Of the nine studies evaluating physical function, all except two showed a significant association.

Interpretation of data concerning function is complicated by the fact that many factors influence the degree of disability expressed by the patient and sometime it may be difficult to prove an association which seems to be obvious. Some of these factors are related to the rheumatoid disease itself and some are unrelated to the underlying inflammatory disease such as age, co-morbidities, psychosocial variables and even measures we use to assess different dimensions of the disease [56]. Disability is influenced by the involvement of both small and large joints however to measure disease activity using DAS28 or 28 joint count methods, we may miss the impact of large joints such as hip which play an important role in physical function [56]. Similarly, joint damage is assessed only in small joints using standard plain radiographs of hands and feet. This does not capture large joint damage or soft tissue abnormalities involving for example tendons or ligaments that may influence function.

Another factor to be considered is the disease duration. In patients with early RA, most of the loss of function is related to inflammation and disease activity, with the potential for improvement with effective treatment. More joint damage with no or limited reversibility occurs with increasing disease duration and this may result in physical function limitation. Therefore a significant component of the loss of function is related to joint damage in established disease [57]. In our review, the impact of disease activity on function was not significant in only two

studies assessing HAQ after a long follow up. None of these analyses was adjusted for joint damage which could potentially have a significant impact on function in these patients who had an established disease at end point. Lindqvist et al analyzed 183 early RA patients' and showed the only significant predictor of HAQ after 10 years was the mean HAQ over the first three months [25]. Similarly, Verstappen et al demonstrated that the baseline functional disability was the only independent factor associated with HAQ after seven years in 112 patients with early RA [27].

Overall, a majority of the studies included in our review reflected a significant effect of disease activity on future function m. None of the reviewed studies assessed the impact of early treatment response or of reaching a desired level of disease activity on future physical function in these patients.

Inconsistency was found for the association between disease activity and joint damage in the prognostic studies reviewed. While the majority of the analyses assessing joint damage as a continuous outcome (absolute joint damage score or score change) found a significant association when the outcome was dichotomized, this association was not observed in most papers. This variation in results may be attributed to certain factors in the analysis.

Distribution histograms of radiographic scores of patients with rheumatoid arthritis are right-skewed. More than 50% of patient have no or minor joint damage and only less than 10% develop high scores [58]. This must be considered when analyzing data with models assuming normality and linearity and appropriate adjustments/transformation should be performed. Majority of our reviewed papers did not discuss this in the methods and it is not clear if it was done. The impact of missing variables on prognostic models and analysis was not well described

or addressed in most studies. Only two imputed the data [14, 39]. A few described patients who were lost to follow-up and compared their baseline characteristics to the ones who completed all assessments [30, 31, 33-36, 40, 42, 45, 46] which in most cases did not show any significant difference. However, missing values are not usually limited to the patients who drop out and may still exist in patients who complete all scheduled follow-ups. Patients with any missing variable would be deleted from the regression models (the most commonly used procedure in our review) which could drop the number of analyzed subjects even further. Only one study reported the actual number of subjects included in the final model [32]. Another factor that could impact the result was confounding effect of treatment variation. Data from observational cohort studies were used in most papers where treatment is not randomized and does not usually follow any specific protocol and this was not addressed in the final multivariate analysis in most cases and could potentially affect the results.

Overall, a significant number of the studies included in our review reflected an association between the disease activity measures and future joint damage progression. All six studies assessing the impact of treatment response (EULAR moderate and good response compare to poor response) or time-integrated disease activity measures showed a significant association with joint damage progression [14-16, 43, 47]. It appears that a significant change in disease activity over time and reaching low disease activity level has a more significant impact on radiographic outcome compare to disease activity level at a single time point (i.e. baseline).

There were a number of causes of between-study variation in included studies influencing interpreting the data. Many of these studies used data from early RA cohorts in different countries. A few used data from clinical trials or extension follow-up of clinical trials with

certain inclusion criteria [14, 24, 26, 27, 36, 37, 40, 53]. Therefore, study populations were heterogeneous in their baseline characteristics. Joint damage was measured by one of the two commonly used methods, Larsen or Sharp van der Heijde score. In the third subgroup of prognostic studies assessing joint damage, the outcome was dichotomized and the cut point was not the same in all. Four studies used Smallest Detectable Difference (SDD) [15, 17, 18, 47], three used Minimal Clinically Important Difference (MCID)[30, 48, 49] and the rest used different definitions (see Appendix 1-C). The underlying heterogeneity in the outcome definition and measurement can contribute to the variation noted in the association of disease activity and joint damage in different studies.

In this review we collected studies evaluating an association between disease activity measures and long-term outcomes. As a measure of disease activity we included any composite measure or joint count score. Although these variables are validated measures of disease activity they are not exactly the same in nature [59, 60]. Composite measures include both patient and physician reported outcomes as well as laboratory test results whereas joint counts are based on physician assessment. This variation in predictors may also contribute to the differences found in the results of prognostic models. Study patients were recruited from mid 1980s to early 2000 and changes in treatment pattern over study years could also impact disease activity, physical function and joint damage.

This review has some limitations. No quality assessment has been performed on the included studies. Given the wide heterogeneity of the articles, the different definitions of the outcome and prognostic factors used, and their different analytical methods, which did not consider the same confounding factors, we were unable to statistically pool the data to perform a meta-analysis. We

simply extracted data from each article and pooled them in tables in a way to give a global idea of the actual literature conclusions on our questions.

Included studies in our review were either exploring the associations between a number of potential prognostic factors and function or joint damage [16, 22-25, 27, 29-33, 35, 40, 42, 43, 48, 49, 52] or were evaluating the independence of the association between a specific prognostic factor and one of the outcome of interest [14, 15, 17, 18, 26, 28, 34, 36-39, 41, 44-47, 50, 51, 53, 55, 61-63]. The impact of disease activity on function or joint damage was not the primary objective of the majority of the included studies and a measure of disease activity was analyzed as a potential confounder/covariate in the multivariate analysis.

To our knowledge this is the first systematic review of prognostic studies assessing the predictive validity of disease activity on functional and radiographic outcome in early RA. We included only studies using multivariate analysis to ensure that our prognostic measures independently predict the main outcomes of interest. We tried to focus on patients with early disease when the irreversible damage is minimal and any improvement in treatment strategy can have a significant impact on patients' future quality of life. Disease activity is a dynamic measure with potentials for significant improvement (remission or LDA), as opposed to other constant prognostic factors such as positive anti-CCP. Our review provides further support for treat to target guidelines that encourage clinicians to focus on this dynamic element as their treatment strategy guide and aim to achieve the goal. We did not include the abstracts that had not been published as a paper because we were aware of the limited data available in the abstracts that may hinder our ability to obtain detailed required information.

In summary, published evidence indicates a link between disease activity and functional outcome or joint damage. However, inconsistency was found for the latter association. There is no prognostic study investigating the impact of reaching low disease activity on functional outcome and only 3 assessed the prognostic role of early treatment response on radiographic damage in early RA patients. Further prognostic studies with large sample size using multivariate analysis are needed to assess the role of early response on long-term outcomes in early RA patients.

3 The impact of reaching low disease activity in the first year of disease on disability or joint damage in patients with early rheumatoid arthritis- results from a multicenter cohort

3.1 Introduction

Rheumatoid arthritis (RA) is characterized by a chronic polyarthritis leading to joint damage and functional disability [64]. Both functional decline and radiographic progression have been strongly associated with continued disease activity [21, 65]. The primary treatment goal in RA is to achieve remission and when not possible low disease activity, as early as possible and to prevent joint damage and excess functional disability[2, 4] .

International treat to target guidelines suggest frequent measurement of disease activity to facilitate achievement of remission early [4]. These recommendations are based on data from strategic trials suggesting that early and aggressive therapy towards a target of remission or low disease activity is associated with improved clinical outcomes. In these trials the primary outcome (disease activity) at end point was compared in patients who were assessed frequently and treated per protocol (tight control arm) with patients who were treated according to the routine practice (usual care arm) as a group [62, 66-68]. None of these studies assessed the functional or radiographic outcomes in patients who achieved the desired disease activity level compared to patients who did not.

The identification of factors indicative of poor outcome early in the course of RA is crucial for tailoring treatment. A number of short and long-term studies have attempted to identify predictive factors of joint damage and disability in RA, but the results are conflicting and there is

no absolute agreement between studies [14-16, 22-24, 28-31, 61]. This could be due to heterogeneous study populations, designs and analyses. Among previous prognostic studies in patients with early RA only a few have evaluated the predictive impact of reaching certain response state (e.g. EULAR good response) on radiographic damage [14-16]. A comprehensive review of the literature showed that no study has assessed the impact of treatment response on functional outcome in early RA (Chapter 2).

Remission is ideal but rare and low disease activity seems like a more realistic goal [69]. The impact of reaching low disease activity on radiographic and functional outcomes in early RA is not well described in the literature. We hypothesized that in patients with early RA reaching low disease activity at one year is associated with improved function (HAQ) and less joint damage in future. In this study, we evaluated thesis hypotheses using data from a large cohort of patients treated mainly with DMARDs in routine practice.

3.2 Study objectives

- To determine whether low disease activity at first year predicts physical disability at 3 years in adult patients with early Rheumatoid arthritis (RA) (symptom duration ≤12 months)

- To determine whether low disease activity at first year predicts joint damage on x-ray at 2 years in adult patients with early RA (symptom duration ≤12 months)

3.3 Methods

3.3.1 Study design

To address the above objectives, data from the Study Of New Onset Rheumatoid Arthritis (SONORA) was studied. SONORA is a prospective observational study involving approximately 98 sites in the United States and Canada. Board certified rheumatologists were invited to participate in this study to enroll and assess patients. Patients who met all inclusion criteria (see 2.2.1) were enrolled in the study and underwent a detailed baseline examination.

During the follow up, patients were sent mail surveys every four months to collect information on medication use, patients reported outcomes, satisfaction with care, and changes in employment status. On an annual basis these surveys were conducted by telephone interview. Clinical exam was performed at the end of year 1 and 2 to assess disease activity as well as laboratory and radiographic changes.

3.3.2 Study population

Patients with signs and symptoms of RA who had been referred to enrolling rheumatologists were included in the study.

3.3.2.1 Inclusion criteria

Patients were eligible to participate in the study if they:

(1) Were aged 18 years or older;

(2) Were new patients presenting to the investigator and were within 3 months from the date of first presentation;

(3) Had at time of enrollment in the opinion of a board certified rheumatologist, signs and symptoms of RA no longer than 12 months; and

(4) Were able and willing to provide informed consent.

3.3.2.2 Exclusion Criteria

Patients would be excluded from the study if they:

(1) Had a diagnosis of chronic juvenile onset RA;

(2) Were using high-dose corticosteroids for a condition other than RA;

(3) Had an underlying acute or chronic disease with high likelihood of dying within six months

3.3.3 Study Plan

3.3.3.1 Physician Based Study Schedule

3.3.3.1.1 Baseline visit

The following were collected at baseline: informed consent, demographics, vital signs, RA and non-RA medical /surgical history, RA therapies and concomitant medications, Physical examination (including 66 Swollen Joint Count (SJC) and 68 Tender Joint Count (TJC)), assessment of pain by patient (VAS 1-10) and global assessment of disease activity by physician and patients (both VAS 1-10), assessment of morning stiffness, biochemistry assessment (including inflammatory markers: C-reactive protein (CRP), Erythrocyte Sedimentation Rate (ESR), auto-antibodies including rheumatoid factor (RF) and anti-CCP antibody), X-ray of hands, physician characteristics and practice settings

3.3.3.1.2 Year 1, Year 2 follow ups

The following were performed at follow ups: vital signs, RA and non-RA medical/surgical History, RA therapy and concomitant medications, physical examination, assessment of pain and disease activity, assessment of morning stiffness, biochemistry assessment, X-ray of hands

3.3.3.2 Patient Based Study Schedule

3.3.3.2.1 Baseline Visit

At baseline the following were collected through a telephone interview: sociodemographic characteristics, diseases activity (assessed by RA disease activity index (RADAI)[70]), current medications, co-morbidities, current employment status, RA and non-RA medical/surgical history, Health Assessment Questionnaire (HAQ), SF-36 health survey, EuroQoL ED-5D Health Questionnaire (EQ-5D), Multidimensional Assessment of Fatigue Scale (MAF), FACIT-Fatigue subscale, arthritis self-Efficacy scales, satisfaction with care and resource utilization.

3.3.3.2.2 Follow-up visits

Similar data (baseline) were collected at year 1, 2, 3 and 4 by telephone interview. Between these telephone interviews, patient reported outcomes were collected through mailed questionnaires every 4 months.

3.3.4 Research Ethics Approval

The SONORA study protocol and informed consent were reviewed by proper Institutional Research Ethics Board at participating sites and were approved.

3.3.5 Data collection

Data were recorded on Case Report Forms and were entered into an electronic database. This database is currently located at TGH research institute and the access is limited to the authorized users.

3.3.6 Main outcomes (dependent variables)

3.3.6.1 Physical function

The HAQ has been used widely in the study of the physical function of normal aging and in many rheumatic diseases. It has become a standard in the assessment of rheumatoid arthritis. Five patient-centered outcome dimensions are conceptualized in the full HAQ including: (1) disability, (2) pain and discomfort, (3) drug toxicity, (4) dollar costs, and (5) death[71].

However, the version that is commonly referred to in the literature is the "short" or "2-page" HAQ. It contains the HAQ Disability Index (HAQ-DI), the HAQ visual analog scale (VAS) for pain, and the VAS patient global health scale. The HAQ-DI assesses a patient's level of functional ability. It evaluates patients through their answers to 20 questions [72]. These questions are organized into 8 categories: dressing, rising, eating, walking, hygiene, reach, grip, and usual activities. Each question is answered on a four level scale of impairment ranging from 0 to 3; 0 = no difficulty; 1 = some difficulty; 2 = much difficulty; and 3 = inability to do.

The final HAQ-DI index ranges from 0 to 3 and is the result of the mean of scores from all eight categories. HAQ-DI scores < 0.3 are considered normal. Scores of 0 to 1 are generally considered to represent mild to moderate difficulty, 1 to 2 moderate to severe disability, and 2 to 3 severe to very severe disability. Average scores that have been reported in a population-based

study are 0.49, and in osteoarthritis and rheumatoid arthritis patients are 0.8 and 1.2, respectively [72].

The HAQ-DI is very responsive to change [72]. The minimal clinically important difference (MCID) in serial HAQ-DI scores has been suggested to be 0.22 [73, 74]. The HAQ-DI has been validated in numerous studies and disciplines [71].

Predictors of higher HAQ found in previous studies include higher baseline HAQ, age, gender (female), longer disease duration, higher tender joint count, higher pain scale and radiographic damage [56, 75-79].

Lindqvist et al showed that the mean HAQ during the first three months was a significant predictor of HAQ at 10 years in their cohort (OR 13.36: 95% CI 5.08 -35.14) [75]. In 1992, Leigh et al demonstrated that the initial disability index is a predominant predictor of HAQ-DI at 8 years along with the following baseline variables: higher age, gender (female), higher tender joint count, pain scale and patient global health assessment [76].

A recent systematic review of prospective studies demonstrated an inconsistency when assessing the association between the baseline radiographic damage and functional disability at the end of follow-up[80]. However, the studies which had performed multivariate regression analysis all showed a significant association. A significant association was found between radiographic progression and functional disability at the end of follow-up and between radiographic progression and change in disability over the follow-up period [79].

3.3.6.2 Radiographic progression

Radiographs provide a measure of permanent damage in RA. Plain radiographs of hands and feet are important in assessing patient with RA over time. The two most widely used measures of radiographic damage are based on the work of Sharp and Larsen [81-85]. The Sharp

method involves separate scores for joint space narrowing and erosion, while the Larsen method is based on a global score of each joint [81].

The original Sharp method [82] included radiographs of the hands and wrists and scored features such as periosteal reaction, cortical thinning, osteoporosis, sclerosis, osteophyte formation, defects, cystic changes, surface erosions, joint space narrowing, and ankylosis [86]. Limitations of several of these features led to their omission from the final score. Therefore, the final Sharp method includes two scores, one for erosions and the other for joint space narrowing [86]. For erosions, 17 areas [5 proximal interphalangeal (PIP); 5 metacarpophalangeal (MCP); 1st metacarpal base (MCB); trapezium and trapezoid as one unit (multiangular); scaphoid, lunate, triquetrum (and pisiform), radius, and ulna bone for each hand and wrist] and for joint space narrowing 18 areas [5 PIP; 5 MCP; carpometacarpal (CMC) 3 to 5; multangularscaphoid, lunate-triquetrum, capitate-scaphoid-lunate, radiocarpal, and radioulnar joints for each hand and wrist] are scored. An erosion score of 0 to 5 was given to each joint according to the number of erosions; "5" represented total destruction. Final erosion score could range from 0 to 170. The final score for joint space narrowing ranges from 0 to 144.

The van der Heijde modification [87, 88] of the Sharp method was designed to overcome the two major limitations of the modified version of Sharp score [81]. Firstly, feet are included. Secondly, some sites which were difficult to see on many radiographs and often were difficult to score leading to inter-observer disagreement are not included. Thus, in the van der Heijde modification of the Sharp method, erosion is assessed in 16 joints [5 MCP, 4 PIP, IP of the thumbs, 1st metacarpal base, radius and ulna bones, trapezium and trapezoid as one unit , scaphoid, lunate] for each hand and wrist and 6 joints (5 MTP, 1st IP) for each foot [86].

In the modified Sharp van der Heijde score (SHS), the maximum erosion score and the maximum joint space narrowing score are 160 and 120 respectively in hands and wrists and 120 and 48 in feet. Therefore, the total Sharp/van der Heijde radiographic score ranges from 0 to 448. [86]. MCID for SHS is roughly 1% of the maximum [81].

In SONORA, hand (and wrist) x-ray was taken at baseline, year 1 and 2 and these were assessed by independent readers (SHS range for hands and wrists= 0-280). Radiographic progression was assessed by score differences between two time points. Reading of one patient's radiographs simultaneously has a major advantage that the rater can correct the score for variation in positioning of the limb or variation of the film quality [89]. When films were grouped per patient, a rater compared all films of one patient and judges whether a change in joint damage had occurred. Progressive disease was defined as a change in the total SHS greater than the smallest detectable change (i.e., 3.4 in this cohort) [89].

Previous studies have shown that higher baseline CRP, higher ESR, positive RF or anti-CCP, smoking history, longer disease duration, higher baseline HAQ, baseline radiographic damage are predictors of future radiographic progression [54, 90-95].

3.3.7 Main predictor (independent variable)

3.3.7.1 LDA measured by Simplified Disease Activity Index (SDAI)

SDAI is a validated measure of disease activity in RA and is the sum of the following variables: Tender Joint Count (28 joints), Swollen Joint Count (28 joints), Patient Global Assessment (0-10), Physician Global Assessment (0-10) and CRP (mg/dl) [96]. Following SDAI cut offs have been proposed and are generally accepted: Remission: SDAI $\leq$3.3; Low disease

activity (LDA): 3.3< SDAI ≤11; Moderate disease activity (MDA): 11< SDAI ≤26; High disease activity (HDA): SDAI >26 [96, 97].

Using available clinical variables collected by treating rheumatologists, SDAI was calculated at three annual time points. The main predictor, being in low disease activity (LDA), was a dichotomous variable ("LDA" yes/no at each time point).

3.3.8 Sample size

If the ratio of LDA to n-LDA (has not reached LDA) patients is 1:4 at 12 months [66], to detect a difference between groups of 0.2 in the HAQ score with 90% power, we would need 335 patients, assuming a standard deviation of 0.45 and a 5% significance level. There were 984 patients enrolled in this cohort.

3.4 Analysis

3.4.1 Statistical software

Analysis was conducted with SAS 9.2 for windows (SAS institute, Inc., Carey, NC). Statistical significance was defined as a p-value < 0.05.

3.4.2 Descriptive statistics

To describe the study population characteristics mean (± standard deviation) for continuous variables and proportions for categorical variable were used. The distribution of continuous variables was evaluated for normality. When the distribution of the variable was skewed, median values (interquartile range) were reported.

3.4.3 Model Building

3.4.3.1 Objective 1: The impact of achieving LDA at 12 months on physical function at 3 years

The main outcome, HAQ-DI at 3 years, was a continuous variable. The main predictor, achieving LDA at 12 months, was a dichotomous variable (LDA y/n). Multiple linear regression analysis was used to compare the outcome between LDA groups with adjustment for covariates selected according to a priori hypothesis based on previous studies or clinical relevance. These include age, gender, baseline medications, anti-CCP antibody status (+/-), RF status (+/-), HAQ and baseline damage (SHS>0 at baseline). Since there were only very few patients on biologic DMARDs at baseline only the use of traditional DMARDs (Methotrexate use defined as y/n) was considered. In addition to the above covariates, an interaction between baseline and 12-months LDA was included to assess whether LDA at 12-months or pattern of LDA over the first year was a better predictor of the outcome.

Multicollinearity occurs when highly correlated independent variables provide redundant information and can affect the parameter estimates. The potential covariates were assessed for existence of multicollinearity using correlation statistics, variance inflation factor (VIF) and Tolerance. In variables with VIF > 2.5 (equivalent to Tolerance < 0.4) multicollinearity would be considered to be significant.

A cohort of patients with available outcome (HAQ at 3 years) and potential predictors was assembled (Complete Cases) and utilized as the analysis cohort in linear models. The number of predictors that could be included in a model was determined by total number of observations/10.

The model assumptions including linearity, normality of residuals and homoscedasticity were checked by normality plots and plots of residuals against independent variables and predicted HAQ (outcome). These plots showed relatively large departure from equal variances. Therefore the model was re-run with log transformation of outcome (HAQ at 3 years transformed to log (HAQ+0.7). We did this since HAQ was 0 for a large number of patients and we therefore added a constant value of 0.7). This improved the pattern of our residual plots. Variables without significant estimates were kept in the model as it was equally important to demonstrate variables which were not significantly associated with the outcome of interest. Standardized estimates were calculated and demonstrated in the model. Coefficient of determination (R^2), which provides an estimate of the amount of variation explained by the model, was used when comparing and selecting final models.

3.4.3.2 Model 2: The impact of achieving LDA at 12 months on x-ray progression at 2 years

The main outcome, developing x-ray progression (ΔSHS> 3.4) at 2 years, was a dichotomous variable and was compared between 12-month LDA groups (LDA vs n-LDA) using multiple logistic regression which was adjusted for covariates selected according to a priori hypothesis based on previous studies or clinical relevance. These include age, gender, disease duration, rheumatoid factor, anti-CCP, baseline medications (DMARDs), baseline LDAS status, HAQ and SHS at baseline. Again, interaction between baseline and 12 month SDAI was included in the model.

Similar to the linear model, a “complete case” cohort was assembled first. The number of variables that could be included in the model was determined by the number in the smallest category of the outcome/10.

Hosmer and Lemeshow goodness of fit statistics was used for assessment of the model fit. To compare the model fit Akaike Information Criterion (AIC) statistics was used. The model with the smallest absolute value of this statistics was considered the better fitting model.

3.4.3.3 Missing data

Analyses of multivariate data are frequently hampered by missing values. The intent of any analysis is to make valid inferences regarding a population of interest. Missing data threatens this goal if it is missing in a way which makes the sample different than the population from which it was drawn. Therefore, it is important to respond to a missing data problem in a manner which reflects the population of inference. In longitudinal studies subjects may drop out early or be unavailable during one or more data collection periods. When collecting questionnaires subjects may be unwilling or unable to respond to some questions. These types of missingness are inevitable, unintended and uncontrolled by the researchers and are one of the main challenges they face in the analysis of their data.

There are various techniques to deal with missing data. Handling missingness by eliminating cases with missing data, also known as “complete case analysis” is the default in most statistical software. This ad hoc technique results in smaller sample size and can also lead to biased results if the remaining cases are not representative of the entire population. Another commonly used method is substitution of missing data with a value (single imputation) such as the mean of the

variable in question. This method reduces the variance artificially and diminishes the relationship with other variables [98].

Even if missing values could be imputed in a way that the distribution and relationship of variables were perfectly preserved the imputed dataset would fail to account for missing data uncertainty. As imputed data are only estimate of actual values, any analysis that ignores this uncertainty will lead to a very small standard error and artificially low p-values and higher rates of type I error [99].

3.4.3.3.1 Mechanism responsible for missing data

Little and Rubin[100] classified missing data mechanisms into three categories:

Missing Completely at Random (MCAR): Here missing cases are not different from non-missing cases. These cases can be deleted from the analysis (complete cases analysis) and the only downside will be loss of power. This rarely occurs in longitudinal studies.

Missing at Random (MAR): Missing cases depend on known values and can be described by variables observed in the data set.

Missing Not at Random (MNAR): Missingness occurs in an unmeasured fashion (non-ignorable or inaccessible) and missing data depend on events or items which the researcher has not measured.

3.4.3.4 Multiple Imputation

In order to have an inferentially useful analysis based on data sets that are partially imputed, two requirements must be met. First, the imputation model must reasonably capture the actual distributional relationships between the unobserved and the observed. Secondly, the analysis must take into account the uncertainty in the imputed values, because no matter how much effort one makes, the imputed values are simply not the real observations [101].

In multiple imputation (MI), missing values for each variable are predicted using existing values from other variables. These imputed values (imputes) with existing data, create "imputed data set".

MI inference involves 3 distinct phases: 1) the missing data are filled in *m* times to generate *m* complete data sets 2) the *m* complete data sets are analyzed by using standard procedures, 3) the results from *m* complete data sets are combined for inference [102].

MI has been shown to produce unbiased parameter estimates when data are MAR or MCAR which reflect the uncertainty associated with estimating missing data. Further, multiple imputation has been shown to be robust to departures from normality assumptions and provides adequate results in the presence of low sample size or high rates of missing data [98].

MI generally assumes that the data are MAR. However it has been shown that the effect of an inaccessible missing data mechanism are often minimal in the implementation of multiple imputation[98]. Figure 3-1 shows three MI steps in SAS:

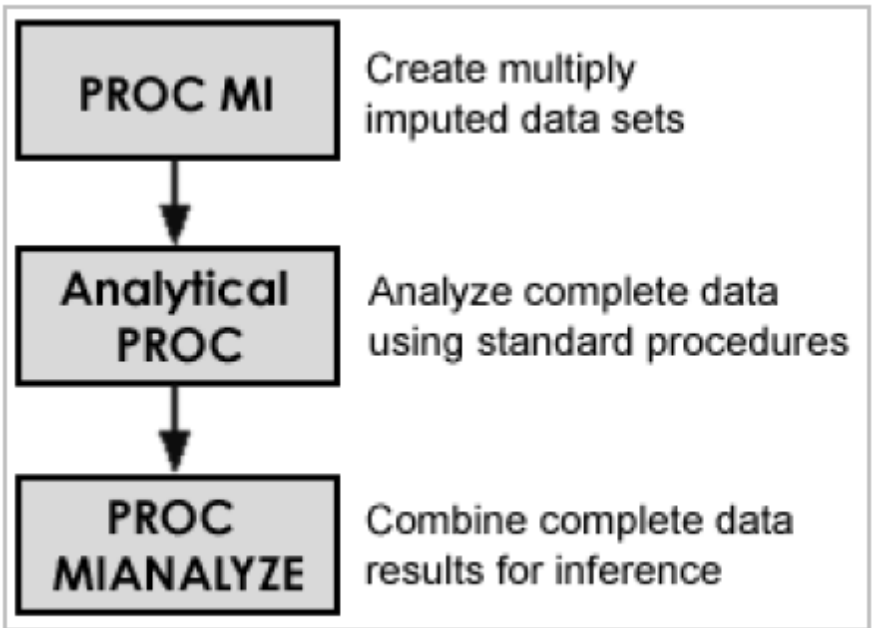

Figure 3-1: The multiple imputation process using SAS software

The MI procedure provides three methods for imputing missing values and the method of choice depends on the type of missing data pattern. A data set with variables Y_1 , Y_2 ,...., Yp has a monotone missing pattern when the event that variables Y_j is missing for a subject implies that all subsequent variables Y_k , $k > j$ are missing for that subject. For data with monotone missing patterns, either a parametric regression method [103] that assumes multivariate normality or nonparametric method that uses propensity score is appropriate [103]. For data sets with arbitrary missing pattern a Markov Chain Monte Carlo (MCMC) that assumes multivariate normality is used to impute all missing values or just enough missing values to make the imputed data sets have monotone missing patterns.

Once the *m* complete data sets are analyzed using standard SAS procedure the MIANALYZE procedure can be used to generate valid statistical inferences about these parameters by combining results from the *m* analyses.

The number of imputed data sets (*m*) is up to the analyst. Commonly, researchers choose between 3-10 data sets [98]. Rubin [103] showed that the efficiency of an estimate based on *m* imputation is approximately :

$$\left(1+\frac{\gamma}{m}\right)^{-1}$$

Where Υ is the fraction of missing information for the quantity being estimated [99]. For example in a data set with 30% missing data the efficiency of MI would change from 94% to 97% if we increase the imputation *m* from 5 to 10. Therefore most analysts use *m*=5.

In the present study, we first identified patients with missing data (focused on outcomes and main predictors). Patients with HAQ available at year 3 (main outcome- model 1) and SDAI at year 1 (main predictor) were compared with patients who did not have HAQ at year 3 or the main predictor with regards to their baseline characteristics to evaluate any significant difference. The same approach was used for radiographic progression (ΔSHS between any two time points i.e. baseline-year 1, year 1-year 2 or baseline-year 2) as the outcome (model 2) and SDAI at year 1 (main predictor) and patients with available outcome and main predictor were compared to patients whose outcome/main predictor were missing with regards to their baseline characteristics (Appendix 2-A).

The pattern of missing data was identified to be arbitrary and MI was conducted using MCMC method. All variables required to build the regression models were included in the imputation model including the outcome. If it was not, the imputed values would not have the same relationship to the dependent variable that the observed values do.

For variables which were dichotomized based on a certain continuous value level (e.g. LDAS, RF status or anti-CCP status) the original value (continuous variable) was imputed and the dichotomous variables were defined and coded again once imputed data sets were established.

The MI procedure assumes that the data are from a multivariate normal distribution when either the regression method or the MCMC method is used. When some variables in a data set are clearly non-normal, it is useful to transform these variables to conform to the normality assumption. With TRANSFORM statement variables are transformed before the imputation process and these transformed variables are displayed in all of the results. By specifying an OUT=option the variable values are reverse-transformed to create the imputed data set. In this study, observed variables analyses had shown a significantly skewed distribution for RF titer, anti-CCP titer, SHS at each time point and SDAI at baseline and year 1 therefore these variables were transformed (logarithmic transformation).

A minimum and maximum limit option was used with PROC MI for each variable to ensure imputed values will remain in acceptable ranges (e.g. maximum SHS=280).

Odds ratios, 95% CI, parameter estimates and corresponding p-values were reported for each model.

3.4.3.5 Exploratory analyses

The impact of reaching LDA, defined by patient reported composite measure of disease activity, RADAI, on functional and radiographic outcomes was explored using the same analysis. RADAI < 2.2 is considered low disease activity level[2].

3.5 Results

3.5.1 Study population

At baseline, of 1101 patients who were considered for enrollment, 984 met the inclusion criteria. Recruitment and cohort maintenance is shown in Appendix 2-B. Of all participants 967 completed the baseline interview (HAQ-DI available). Baseline x-ray was obtained in 735 patients. Total sharp score was recorded for 683 patients and 52 patients had either total joint space narrowing (TJSN) or total erosion (TERO) scores missing. Thirty two patients had one or more disease activity measures missing and SDAI could be calculated for 952 patients.

SDAI was available for 752 patients at year 1 and for 699 patients at year 2. Overall, 212 (21%) patients did not have any x-rays at baseline or during the follow-up and in 59 patients only one of TERO or TJSN was recorded therefore the total Sharp score could not be evaluated. Radiographic scores (SHS) were available for 685 and 574 participants at Year 1 and Year 2, respectively.

Patient interviews were conducted for the majority of patients on a yearly basis. HAQ-DI was available for 899, 846 and 801 patients in year 1, 2 and 3 respectively.

3.5.2 Baseline characteristics

Total cohort's baseline characteristics are demonstrated in Table 3-1. Mean age was 53 years (± 14.8) and 72% were women. Anti-CCP titer was available in 774 patients and more than half (53.4%) were positive. Medication data was available for 983 patients and showed most cases were treated with DMARDs at baseline (74%). Glucocorticoidsteroids (oral or parenteral)

Table 3-1 Baseline characteristics

Variables	Total cohort (n=984)
Female n (%)	708 (72)
Age (yr) *	53 (14.8)
Disease duration (days)*	157 (95)
RF positive (n=975) n (%) **	593 (61)
Anti-CCP positive (n=774) n (%)‡	413 (53.0)
SDAI *†	30.5 (16.6) [28.1, 17.8-41.8]
SJC *†	9.4 (7.1) [8.0, 3.0-14.0]
TJC*†	10.1 (8.0) [9.0, 3.0-16.0]
MD global [0-10]*†	4.8 (2.1) [5.0, 3.0-6.0]
Patient global [0-10]*†	4.7 (2.4) [5.0, 3.0-7.0]
CRP mg/dl *†	1.4 (1.5) [0.8, 0.8-1.1]
SHS *†	5.03 (7.3) [3.0, 0.0-7.0]
DAS28*†	5.02 (7.3) [3.0, 3.9-5.9]
RADAI*†	4.32 (1.9) [4.28, 2.9-5.6]
HAQ *†	1.0 (0.72) [1.0, 0.4-1.6]
DMARDs n (%)	820 (86%)
Biologics n (%)	18 (2%)
GCS n (%)	641 (67%)

*mean (SD), ** RF >20 was considered positive, ‡ anti-CCP > 20 units was considered positive, †median [IQR], SHS=Sharp van der Heijde Score

were used in 54% of patients and the majority received them in combination with DMARDs. Only 2% were treated with biologic-DMARDs (anti-TNF agents or Anakinra).

3.5.3 Description of main predictor and outcomes

Disease activity (SDAI), HAQ and SHS at baseline and follow ups were not normally distributed. Tests for non-normality were significant for these variables ($p<0.05$, Appendix 2-C). Disease activity, radiographic and functional outcomes at year 1, year 2 and year 3 follow ups are demonstrated in Table 3-2.

Table 3-2 Disease activity (SDAI), HAQ and Sharp Score at different time points

Variable	Baseline	12 mo	24 mo	36 mo
SDAI mean (SD) median (IQR)	**30.5** (16.6) 28.1 (18.0-42.0)	**18.8** (14.3) 14.5 (8.0-27.0)	**16.6** (13.6) 12.8 (7.0-22.0)	NA
HAQ mean (SD) median (IQR)	**1.00** (0.72) 1.0 (0.4-1.6)	**0.82** (0.71) 0.75 (0.1-1.3)	**0.77** (0.72) 0.63 (0.1-1.3)	**0.7** (0.70) 0.6 (0.1-1.3)
SHS mean (SD) median (IQR)	**5.03** (7.31) 3.0 (0.0-7.0)	**6.21** (8.76) 3.0 (1.0-8.0)	**6.39** (9.24) 3.0 (1.0-8.0)	NA

SDAI=Simplified Disease Activity Index, SHS: Sharp van der Heijde Score, NA=not available

Functional disability improved over time (Figure 3- 2). At year 1, 296 (30%) achieved LDA. RADAI also improved over time from 4.32 (1.91) (mean(sd)) at baseline to 2.80 (2.18) at year 2 (Appendix 2-D).

Of 683 patients who had x-ray at baseline, 503 (74%) patients had damage (SHS>0). Of these, 282 (41%) had significant damage (SHS $\geq$ 3.4) at baseline. Over the 2 year follow-up radiographic progression was observed in 116 (17%) patients and in the majority (76 patients) progression had occurred between the baseline and year 1. At baseline 216 (22%) had no disability and the rest had some degree of functional limitation (Table 3-3).

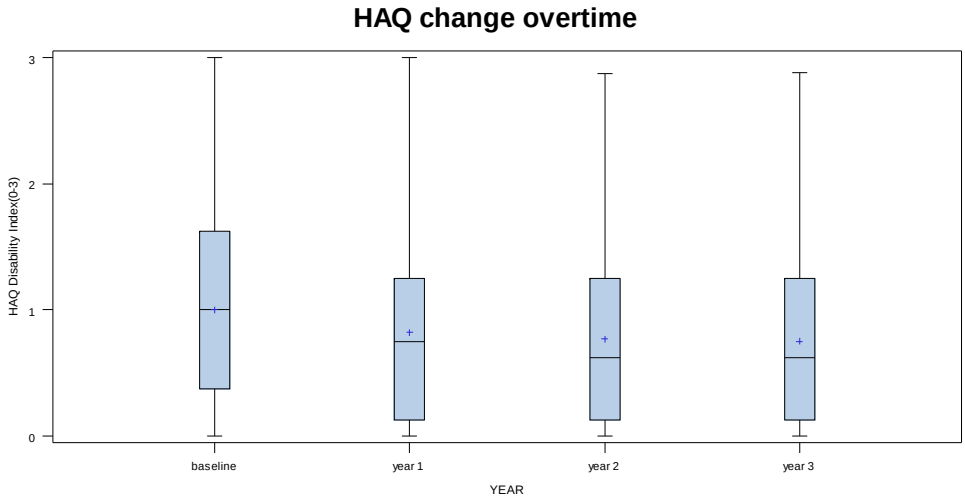

Figure 3-2 HAQ pattern over 3 year follow up

Radiographic and functional outcomes are compared in year 1 LDA and n-LDA groups in Table 3-4. Patients who achieved LDA had lower HAQ at year 3 and radiographic progression rate was lower in these patients although the difference did not reach a statistical significance (Table 3-4).

Table 3-3 Disability severity at baseline

HAQ disability index	Baseline
Not disable (HAQ < 0.3)	216 (22%)
Mild (0.3≤ HAQ <1.0)	262 (27%)
Moderate (1.0≤ HAQ <2.0)	372 (39%)
Severe (HAQ ≥ 2.0)	117 (12%)

Table 3-4 Comparison of functional and radiographic outcomes in two groups

	Achieved LDA at yr 1 (LDA)	Did not achieve LDA at yr 1 (n-LDA)	p-value
HAQ at year 3, mean (sd)	0.48(0.58)	0.93(0.71)	<.0001
X-ray progressed, n (%)	34(11.5%)	78(16%)	0.08

3.5.4 Achieving LDA at year 1 and functional outcome

Linear regression models included variables of interest based on existing evidence and clinical rational including age, gender, baseline HAQ, RF status, anti-CCP status, baseline SHS and MTX use. The main predictor, being in LDA at year 1 was included in addition to interaction between the LDA status at baseline and year 1. Sharp score components, total erosion (TERO) and total joint space narrowing (TJSN) scores, were considered as covariates as well. Univariate regression analysis results are shown in Table 3-5.

Test of Multicollinearity was performed (Table 3-6). As VIF and Tolerance were borderline, Tetrachoric correlation was assessed between RF and anti-CCP and that was significant (r=0.89, p=0.02). Therefore these two variables were not included in the same model.

3.5.4.1 Complete case analysis

Based on covariates, complete case analysis was performed on 436 patients with available data for all potential confounders and outcomes. “LDA yr 1 x LDA baseline” interaction was included in the first model. This interaction was not significant and was removed.

Table 3-5 Univariate regression analysis, Outcome= HAQ year 3 (n=984)

Predictors	β coefficients	P-value	N=984
LDAS yr 1	-0.3828	<.0001	696
LDAS baseline	-0.3183	<.0001	801
Age (yr)	0.0101	<.0001	801
Gender (female)	0.2281	<.0001	801
RF positive	0.0382	0.41	796
Anti-CCP +	0.0065	0.91	641
MTX use baseline	0.0667	0.17	801
Sharp score baseline	0.0191	<.0001	603
TJSN score baseline	0.0429	<.0001	642
TERO score baseline	0.0184	0.0013	625

TJSN=Total Joint Space Narrowing; TERO=Total Erosion;

As described above (see 1.9.3.1) the outcome, HAQ at year 3, was transformed to log (HAQ+0.7) to improve the residual plots’ pattern (Appendix2- F). To assess the effect of joint damage measures, total sharp score or TJSN and/or TERO were included in each model along with the rest of covariates listed above.

Models including total sharp score or TERO alone had the lowest R^2 (Table 1-Appendix 2-G). Models including both TJSN and TERO are shown in Table 3-7.

Achieving LDA at year 1 was significant in all models. Lower baseline HAQ, LDA at year 1, younger age and gender (male) were associated with improved functional outcome at year 3 (Table 3-7).

Table 3-6 Test for Multicollinearity (Primary outcome: Functional disability)

	VIF	Tolerance
LDAS yr 1	1.13297	0.88264
LDAS baseline	1.07873	0.92701
HAQ baseline	1.20071	0.83284
Female	1.12007	0.89280
Age	1.23958	0.80673
RF	1.91057	0.52340
Anti-CCP	1.95372	0.51184
MTX	1.05980	0.94358
CRP baseline	1.04498	0.95696
TJSN baseline	1.18473	0.84407
TERO baseline	1.20819	0.82768

TJSN=Total Joint Space Narrowing; TERO=Total Erosion; VIF= Variance Inflation Factor

Table 3-7 Linear regression models. Complete case analysis after transformation [outcome = log (HAQ + 0.7)]

	Model 1 with anti-CCP		Model 2 with RF	
	β coefficient (95% CI)	**p-value**	**β coefficient (95% CI)**	**p-value**
LDAS1	-0.1298 (-0.2094,-0.0503)	0.0014	-0.1296 (-0.2095, -0.0498)	0.001
LDAS0	0.0258 (-0.0903, 0.1417)	0.66	0.0285 (-0.0877, 0.1449)	0.63
Age	0.0050 (0.0020, 0.0079)	0.001	0.0047 (0.0020,0.0077)	0.002
Female	0.0926 (0.0032,0.1818)	0.04	0.0876 (-0.0014,0.1767)	0.05
HAQ (BSL)	0.3374 (0.2802, 0.3946)	<0.0001	0.3391 (0.2819,0.3964)	<0.0001
TJSN (BSL)	0.0107 (-0.0015, 0.0228)	0.085	0.0104 (-0.0018, 0.0225)	0.09
TERO (BSL)	-0.0014 (-0.0103, 0.0075)	0.76	-0.0010 (-0.0099, 0.0079)	0.82
Anti-CCP	0.0442 (-0.0307, 0.1190)	0.25		
RF			0.01280 (-0.0624,0.0881)	0.74
MTX	0.0040 (-0.0704,0.0784)	0.92	0.0082 (-0.0661,0.0824)	0.82

Model 1: R^2=0. 362; Model 2: R^2=0. 360

Total sharp score and TERO score were not associated with HAQ. Higher TJSN was associated with higher HAQ but this association did not reach the statistical significance (p=0.09). Baseline HAQ appeared to be the strongest predictor.

3.5.4.2 Analysis of the imputed data

The result of the analysis after imputation is shown in Table 3-8. Overall, significant predictors in complete case analysis remained significant after the imputation. To demonstrate the impact of TJSN and TERO when the model was adjusted for both, models including both of these variables are shown (Table 3-8).

Models including each one of these two sharp score components and a model with total sharp score as a covariate, were analyzed as well (Table 2. Appendix 2-G). TJSN and total sharp score but not TERO remained significant in those models. The R-square is reported for each imputation in Model 1 and 2 (Table 3-8).

Table 3-8: Linear regression models. Analysis after Multiple Imputations (n=984) outcome: HAQ at year 3

	Model 1 with anti-CCP		**Model 2 with RF**	
	β coefficient (95% CI)	**p-value**	**β coefficient (95% CI)**	**p-value**
LDAS1	-0.2121 (-0.3049, -0.1193)	<.0001	-0.2105 (-0.3032, -0.1176)	<.0001
LDAS0	0.0339 (-0.1040, 0.1719)	0.62	0.0332 (-0.1049, 0.1713)	0.63
Age	0.0050 (0.0020, 0.0080)	0.002	0.0050 (0.0020,0.0081)	0.001
Female	0.0956 (0.0146,0.1767)	0.02	0.0960 (0.0147,0.1772)	0.02
HAQ (BSL)	0.4707 (0.4168, 0.5247)	<.0001	0.4704 (0.4165, 0.5244)	<.0001
TJSN (BSL)	0.0155 (-0.0010, 0.0320)	0.06	0.0156 (-0.0008, 0.0319)	0.06
TERO (BSL)	0.0040 (-0.0090, 0.0169)	0.51	0.0039 (-0.0091, 0.0167)	0.52
Anti-CCP	0.0022 (-0.0789, 0.0832)	0.95		
RF			0.0320 (-0.0391,0.1032)	0.38
MTX	-0.0209 (-0.0953,0.0536)	0.58	-0.0234 (-0.0966, 0.0497)	0.53

Model 1 R^2 (imputation 1-5): 0.359, 0.391, 0.378, 0.384, and 0.378, average=0.378
Model 2 R^2 (imputation 1-5): 0.359, 0.392, 0.379, 0.385, and 0.379; average=0.379

3.5.5 Achieving LDA at year 1 and radiographic damage

Logistic regression models assessed variables of interest based on existing evidence and clinical rational including age, gender, baseline HAQ, CRP, RF status, anti-CCP status, baseline

radiographic damage (SHS>0) and MTX use at baseline. The main predictor, LDA at year 1 in addition to LDA status at baseline and year 1-baseline LDA interaction were included in the model. Univariate regression analysis results are shown in Table 3-9.

Table 3- 9 Univariate regression analysis. Outcome: X-ray progression at 2 years

Predictors	OR (95% CI)	p-value	n=984
LDA year 1	0.62 (0.40-0.97)	0.03	654
LDA baseline	0.48 (0.21-1.07)	0.07	676
Age	1.02 (1.01-1.03)	0.01	694
Gender (female)	0.79 (0.51-1.23)	0.30	694
RF +	2.1 6 (1.38-3.40)	0.0008	689
Anti-CCP +	2.82 (1.69-4.71)	<0.0001	543
CRP baseline	1.24 (1.10-1.40)	0.0005	685
SHS BSL >0	3.08 (1.71-5.54)	0.0002	675
SHS baseline	1.08 (1.05-1.11)	<.0001	675
HAQ-DI baseline	1.23 (0.93-1.62)	0.15	692
MTX use baseline	1.19 (0.80-1.77)	0.40	694

Test of Multicollinearity was performed (Appendix 2- E). As VIF and Tolerance were borderline for RF and anti-CCP, tetrachoric correlation was assessed between these two variables and that was significant (r=0.89, p=0.02). Therefore these two variables were not included in the same model.

3.5.5.1 Complete case analysis

Based on variables of interest, complete case analysis was performed on 485 patients with available data for all potential confounders and outcomes. "LDA yr 1 x LDA baseline" interaction was included in the first model. This interaction was not significant and was removed. Among these 485 patients 80 had progressed radiographically over two years. It would limit the number of variables in each model to 8 (see 3.4.3.2). Hosmer and Lemeshow goodness of fit statistics showed a non-significant p-value indicating those models fit the data except one (Table 3-10). Models with lower AIC and higher c-statistics including RF and anti-CCP are shown in Table 3-10 (see Appendix 2-H for all models).

Table 3-10 Logistic regression models. Outcome: 2 year x-ray progression. Complete case analysis (n=485)

	Model 1 with anti-CCP			**Model 2 with RF**		
	OR	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDAS1	0.74	0.43-1.30	0.29	0.79	0.45-1.36	0.39
LDAS0	0.48	0.18-1.29	0.15	0.50	0.19-1.34	0.17
Age	1.01	0.99-1.03	0.31	1.01	0.99-1.03	0.44
Female	0.60	0.35-1.03	0.06	0.55	0.33-0.94	0.03
CRP BSL	1.13	0.91-1.41	0.27	1.12	0.91-1.39	0.29
SHS BSL>0	2.73	1.27-5.88	0.01	2.69	1.26-5.77	0.01
MTX	0.84	0.50-1.41	0.50	0.94	0.56-1.56	0.80
Anti-CCP	3.54	2.01-6.25	<.0001	-	-	-
RF	-	-	-	2.32	1.32-4.08	0.003
c-statistics	0.736			0.710		
AIC	407.06			419.02*		

*Hosmer and Lemeshow Goodness-of-Fit Test p=0.03;

Higher baseline total sharp score and positive anti-body status were both associated with radiographic progression.

3.5.5.2 Analysis of the imputed data

The result of the analysis after imputation for models including same variables used in complete case analysis is shown in Table 3-11 (see Appendix H Table 2 for all models). LDA at year1 was associated with less radiographic progression. RF, anti-CCP antibody and baseline damage remained as the strongest predictors of joint damage. Older age and higher baseline CRP were associated with damage progression over 2 years.

Table 3-11 Logistic regression models. Outcome: x-ray progression at year 2. Imputed data analysis

	Model 1 with anti-CCP			Model 2 with RF		
Predictor	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDAS1	0.61	0.38-0.98	0.04	0.64	0.39-1.39	0.07
LDAS0	1.02	0.41-2.52	0.96	1.04	0.42-2.57	0.92
Age	1.02	1.01-1.04	0.008	1.02	1.01-1.04	0.004
Female	0.86	0.54-1.36	0.51	0.82	0.52-1.29	0.38
CRP BSL	1.15	1.00-1.32	0.05	1.17	1.02-1.34	0.03
SHS BSL>0	3.38	1.85-6.18	<.0001	3.32	1.82-6.07	0.0001
HAQ BSL	1.04	0.81-1.35	0.74	1.05	0.81-1.36	0.72
Anti-CCP	2.18	1.54-3.07	<.0001	-	-	-
RF	-	-	-	1.90	1.19-3.01	0.009
c-statistics range†	0.709- 0.725			0.705- 0.720		
AIC range†	982.07- 1035.33			997.56-1050.44		

† shows statistic range for 5 imputed datasets

3.5.6 Exploratory analysis

3.5.6.1 Association between patient reported disease activity measure (RADAI) and x-ray progression

Logistic regression was used to assess this association. Covariates were similar to the ones included in the previous analyses (i.e. age, gender, baseline HAQ, CRP, RF status, anti-CCP status, baseline radiographic damage (SHS>0) and MTX use). RADAI was available at 4, 8 and 12 months. In order to assess the impact of very early response, i.e. low disease activity before 6 months, two sets of logistic regression analysis were considered. The main predictor for one set was achieving LDA (RADAI< 2.2) at 4 months and for the other was LDA at 12 months. Same as previous analysis, year 1-baseline LDA and 4-month interactions were included in each model. These were not significant and were removed. The outcome for both was x-ray progression at 2 years. Complete case cohort was first identified. To make complete case analysis models comparable, cases with all covariates, 4 months and year 1 RADAI available were considered complete and used in the model (n=479). X-ray progression had occurred in 77 patients and this would limit the number of included variables in the model to 7.

LDA (RADAI<2.2) at 4 months was associated with less radiographic progression in most models but this association was not observed between LDA at 12 months and radiographic damage. In all models, positive autoantibody status and baseline damage were associated with damage progression (see Appendix 2-I).

In the analysis of imputed data, positive autoantibodies (RF and anti-CCP) and baseline damage were significant predictors of damage. The association between LDA at 4 or 12 months and

radiographic progression was not significant except for LDA at 12 months in 2 models (see Appendix 2- J).

3.5.6.2 Association between patient reported disease activity measure (RADAI) and function (HAQ)

Multivariate linear regression was used to assess this association. Covariates were similar to the ones included in the previous analysis where SDAI was used as the disease activity measure (i.e. age, gender, baseline HAQ, RF status, anti-CCP status, baseline radiographic scores , baseline HAQ and MTX use). RADAI was available at baseline, 4, 8 and 12 months. In addition to the LDA at year 1, to assess the impact of very early response, i.e. achieving low disease activity before 6 months, on function, two sets of linear regression analysis were considered. The main predictor for one set was achieving LDA (RADAI< 2.2) at 4 months and for the other was LDA at 12 months. Same as previous analysis, year 1-baseline LDA and 4 month-baseline interactions were included in each model. These were not significant and were removed. The outcome for both sets was HAQ at year 3. Complete case cohort was first identified. To make complete case analysis models comparable, cases with available baseline,4 months, year 1 RADAI and all other covariates were considered complete and used in the model (n=449).

LDA (RADAI<2.2) at 4 months and 12 months were associated with less functional disability in complete cases analysis. Age and baseline HAQ were the other two significant predictors. Older age and higher HAQ were associated with higher HAQ at 3 years in both sets. R^2 was higher for models including 12 month HAQ (see Appendix 2-K).

In the analysis of imputed data, LDA, baseline HAQ and age remained significant in both 4 month and 12 month sets (see Appendix 2-L). Female gender and higher baseline TJSN were two additional predictors of 3 year disability.

3.6 Discussion

3.6.1 *Key findings and clinical implications*

The current study demonstrated the impact of reaching low level of disease activity on two main outcomes in patients with early RA. It showed patients who reach LDA at one year are less likely to show functional deterioration at 3 years. This association was significant in both complete case analysis and after multiple imputations. Our exploratory analysis showed that this finding was persistent when LDA was assessed by both disease activity measures (SDAI and RADAI). Using RADAI, it was shown that reaching LDA as early as 4 month predicts lower HAQ at end point. Other significant predictors of disability were higher HAQ, older age and gender (female). Total sharp score (SHS) at baseline did not show any significant association with HAQ at 3 years however between two major components of this score, less joint space narrowing was found to be associated with better functional score at end point.

The association between LDA at year 1 and 2-year radiographic progression was found to be significant in imputed data analysis. It was demonstrated that achieving low disease activity (SDAI <11) predicts less radiographic damage progression. Other significant baseline predictors of radiographic damage progression over 2 years were positive autoantibodies (RF or anti-CCP), radiographic damage, older age and higher CRP. The exploratory analysis of the impact of reaching LDA assessed by RADAI at 4 months and 12 months on joint damage revealed that

achieving RADAI < 2.2 at 4 months was associated with less damage in complete case analysis. This association was not statistically significant for achieving low RADAI at 12 months in complete case analysis and for neither time points in most models in imputed data analysis, however the direction of change was in favor of a protective impact especially for low RADAI at 12 months which was significant in two models. Positive autoantibodies and baseline joint damage remained as significant predictors of x-ray progression. Higher CRP and older age were shown to be associated with more damage progression in imputed data analysis.

Our findings support current early treat-to-target recommendations [2, 4]. It consistently showed that reaching LDA at one year is associated with lower HAQ at 3 years. Our exploratory analysis confirmed this finding and showed achieving LDA as early as 4 months (by RADAI criteria) predicts less functional deterioration. Demonstrating a clear association between disease activity, functional disability and joint damage over time is challenging. In recent years, several investigators have been trying to identify this complex relationship. It is generally believed that patients' functional capacity is mainly influenced by inflammation and disease activity during the early stages of disease which will improve with treatment. But with increasing disease duration patients develop damage resulting in functional decline due to this irreversible phenomenon more than disease activity which is usually not as severe as early phases [57, 104, 105]. In one study, RA patients with disease durations < 4 years were followed over 10 years, and using Generalized Estimating Equations (GEE) a longitudinal association between SHS, HAQ and grip strength was demonstrated. They considered ESR as the measure of disease activity [106]. An increase of 10 units in the SHS was associated with a 0.03 unit increase in HAQ score. In 2011, Smolen et al. estimated the numerical value for damage-related change in the physical function deterioration (DAM-HAQ) in clinical trials to be 0.01 points per unit

change in SHS[105]. In our study, evaluation of our first objective, demonstrated that the baseline SHS was associated with patients' function at end point. Further assessment of this association showed that in fact it was the joint space narrowing and not the joint erosion which predicted patients' function at 3 years. This finding is in line with the results of Aletaha et al study which concluded that cartilage damage is more clearly associated with irreversible functional limitation compare to bone damage [107] .

Our results showed that in early RA patients there was a significant association between physical function at 3 years and disease activity at an earlier (4 months or 1 year) time point. When the baseline disease activity status was included in the model, it was still year-1 LDA (or 4 month LDA) that had a strong association with functional outcome indicating that a response criteria (i.e. reaching LDA) is a better predictor of this outcome compare to LDA at baseline. These findings suggest that clinicians should make their best effort to improve patients disease activity as early as possible as it is associated with better functional capacity in future. There are certainly many other unknown confounders that are not captured in routine cohort studies or clinical trials such as depression, soft tissue injuries (e.g. tendon rupture) or co-morbidities. These factors may influence physical function and are not measured routinely and not considered in the analysis but can significantly affect the outcome and make it sometime difficult to demonstrate a clear association. These results were consistent with previous studies identified in our literature review [21, 23, 24, 26]. Only two studies failed to show a significant association between disease activity measures and function in their multivariate analysis [25, 27]. The follow-up duration were 7 [27]and 10 years [25]in these studies. Lack of association could be due to the long follow-up duration where the role of joint damage would be more prominent compare to the disease activity.

We also showed that achieving LDA at year 1 reduces the risk of radiographic progression. The baseline predictors of joint damage in early RA have been assessed in a number of studies which have shown conflicting findings [14-18, 28, 30, 33-40, 42-51, 53, 61, 62]. Our literature review demonstrated that about a third of included studies showed lower disease activity, measured by composite scores or joint counts, was associated with less radiographic damage [14-18, 21, 31, 38, 43, 47, 53]. However majority did not show any significant association [22, 27-30, 33-37, 39, 40, 42, 44-46, 48-51]. As discussed earlier (see Chapter 2), differences in patient populations, methods and analysis have likely contributed to this heterogeneity. The predictive value of disease activity (at baseline or during the follow up) on damage has very occasionally been the primary objective of prognostic studies in early RA. Achieving a desired disease activity status has been assessed only in 3 studies [14-16] and showed EULAR good response at 6 months was associated with less damage when compared to no-response. Baseline damage, positive anti-body status (both anti-CCP and RF) and high CRP were significant predictors of radiographic damage progression in our cohort and this is in line with previous studies' findings.

The association between LDA and damage progression was not statistically significant in all models (models included RF) although the trend was towards a preventive effect which was near significant ($p=0.06$ and $p=0.07$). In our study x-ray progression had occurred between baseline and year 1 in most patients who progressed over 2 years, i.e. prior to our main predictor (LDA at year 1) and that could potentially result in a less significant association between year 1 LDA and the outcome. However, despite occasional flares, disease activity usually follows a steady trend in most patients and the LDA status at year 1 very likely represents how well patients' disease had been controlled within a reasonable interval prior to or after that assessment. We also kept

the baseline LDA status and its interaction with year 1 LDA in the model to assess whether LDA at 12-months or pattern of LDA over the first year was a better predictor of the outcome.

Although patient reported LDA (RADAI <2.2), either at 4 months or at 12 months, was associated with less radiographic damage this association did not reach the statistical significance in most models. RADAI is a valid measure of disease activity in RA patients and has been shown to have a significant correlation with other composite measures [108]. As noted in methods, in addition to the clinical parameters, CRP is included in SDAI which may contribute to its more significant predictive value for joint damage progression.

Overall multiple imputation (MI) improved out study power although it did not change the direction of associations for many variables. Initially there seemed to be no selective drop-out, suggesting that the available patients were a good representation of the total population but due to the large number of incomplete cases who were eliminated from multivariate analyses it seems that those deleted patient were likely different in a number of predictors and the complete case analysis could result in biased results. Longitudinal studies are a valuable resource for prognostic studies and MI is a practical solution for management of missing cases that should not be under used.

3.6.2 Study weaknesses

Our study has certain limitations. In SONORA only hand/wrist x-ray was obtained. Standard imaging in RA patients include hands/writs and feet. It is possible that more damaged joints could be identified if foot images were taken and that would improve our study power.

Missing data were one of our main challenges. Lost to follow-up was relatively significant for clinical data (lost to follow-up for patient reported data was sparse) but the main source of missing data was missing variables in incomplete cases that were deleted from the multivariate regression analyses. This had reduced the number of patients included in the final models to half.

Current guidelines recommend early/aggressive therapy aiming towards achieving remission/LDA by 6 months. Our first clinical assessment (i.e. SDAI) was at 12 months and it would be ideal to have earlier clinical follow-upmeasures of disease activity. However, we were able to analyze a patient reported disease activity measure at an early time point (RADAI at 4 months) but as discussed above the predictive impact of this measure was not identical to SDAI.

3.6.3 Study strengths

In this analysis we have used SONORA database which is comprised of a large number of patients (n=984) and represents a typical real-world early RA cohort. X-rays were available for majority of (about 70%) of included patients and were scores according to a standard scoring system by experts with satisfactory inter-observer and intra-observer agreement. This is a unique value of SONORA as it is quite challenging to obtain, standardize and transfer images for reading in a multicenter (98 centers) observational study which recruits patients from different settings (both community and academic). Radiographic progression occurs early in the disease process [109]therefore the follow-up duration (2 years) was reasonable.

Missing data are one of the main limitations in observational studies. Our analysis was not an exception however we tried to address this issue, first by confirming that there was no selective drop-out, suggesting that the available patients were a good representation of total population and then by running both complete case and imputed data analysis after performing MI. In this

method missing data is accounted by restoring not only the natural variability in the missing data, but also by incorporating the uncertainty caused by estimating missing data. Therefore it is intended to preserve important characteristics of the data set as a whole. MI is a powerful and widely accepted method for management of missing data but it is not implemented by many investigators, likely because of lack of familiarity. Longitudinal data from observational studies are a valuable resource for prognostic studies but usually suffer from missing values and MI is a practical solution for management of missing cases that should not be underutilized. It should be noted that the drop-out rate was quite low for patient reported outcome assessments (interviews/surveys) in this large cohort which adds to the data quality.

In our exploratory analysis, we were able to assess the predictive impact of RADAI on damage and disability. To our knowledge this is the only study that has evaluated RADAI as the key predictor of joint damage or disability in RA. This variable was available through surveys sent to patients every 4 months and this enabled us to explore and compare the above associations at earlier time points.

3.6.4 Unanswered questions and future research

This study was focused on two main RA outcomes, joint damage and disability, however the impact of response to treatment on patients' long-term quality of life, work productivity and resource utilization remains to be evaluated. Longitudinal data from observational cohorts such as SONORA, provide valuable resource for this type of investigations.

Several studies have been designed to assess predictors of long-term outcomes in RA and have identified factors associated with outcomes at a group level. Demonstration of negative impact of poor prognostic factors is extremely valuable and essential however quantification of the risk

associated with various factors in each patient scenario is more likely practical for clinicians who see these patients everyday and decide about treatment based on their background knowledge, practice guidelines and patient preferences. In recent years, several groups have been working on prediction matrices for x-ray progression in RA [63, 110-112]. Most of these matrices are developed on clinical trial patients that are not representative of real-world practice. In 2009, a prediction matrix for x-ray progression was developed in SONORA patients[113] and a more recent study presented a risk prediction model for rapid radiographic progression in another early RA cohort (ESPOIR) [114]. Risk stratification models seem to be the main theme for future prognostic studies in RA and appear to be more practical and clinically meaningful for practitioner who may use them routinely. These can help researchers identify risk of most significant outcomes for RA at an individual patient level.

Prognostic studies are investigations of future events or assessment of associations between risk factors and health outcomes. [115]. Hayden et al identified three phases of explanatory prognosis investigation: Phase 1 studies identify associations between a number of potential prognostic factors and a health outcome (hypothesis generating evidence) and are the most common phase of prognostic investigation. Phase 2 studies aim to measure the independent effect of a prognostic factor while controlling for confounders and Phase 3 studies attempt to describe the complexity of the prognostic pathways or processes [116]. It appears that over the last 2 decades prognostic studies in RA have focused on investigating association between risk factors and certain outcomes including joint damage or disability. A few have investigated the independent association between a specific factor and these outcomes including our analysis. Now that several risk factors have been identified in RA, there is an opportunity for phase 3 prognostic investigations that attempt to apply knowledge from the previous phases on

independent associations and incorporate other knowledge and expert opinion in this field to develop theoretical frameworks which can provide evidence for the mechanism of action of prognostic factors on the outcome and help understand their complex relationships.

4 Conclusions

In this thesis, we first collected published evidence investigating association between disease activity and long-term outcomes in early RA. Our review showed that published evidence supports an association between disease activity and future disability or joint damage. We also demonstrated that there was no prognostic study evaluating association between reaching low disease activity and disability and there were very few studies addressing association between early good clinical response and less radiographic progression in this patient population.

We then assessed the predictive impact of reaching LDA on disability and damage in a large observational cohort of early RA patients with prognostic models using multivariate regression analyses. Our results showed a significant association between LDA (based on SDAI or RADAI) and both outcomes. To our knowledge this is the first study that evaluated the prognostic value of a patient reported disease activity measure, RADAI, on long-term outcomes in patients with early RA.

Our findings support current treat to target guidelines recommending remission or LDA as the goal of therapy. It also shed further light on the impact of patient driven measures of disease activity, which are becoming an integral aspect of disease assessment in RA, on future outcomes in patients with early disease[117].

References:

1. Bombardier, C., et al., *Canadian Rheumatology Association recommendations for the pharmacological management of rheumatoid arthritis with traditional and biologic disease-modifying antirheumatic drugs: part II safety.* J Rheumatol, 2012. **39**(8): p. 1583-602.

2. Bykerk, V.P., et al., *Canadian Rheumatology Association recommendations for pharmacological management of rheumatoid arthritis with traditional and biologic disease-modifying antirheumatic drugs.* J Rheumatol, 2012. **39**(8): p. 1559-82.

3. Singh, J.A., et al., *2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis.* Arthritis Care Res (Hoboken), 2012. **64**(5): p. 625-39.

4. Smolen, J.S., et al., *Treating rheumatoid arthritis to target: recommendations of an international task force. .* Ann Rheum Dis, 2010. **69**: p. 631-7.

5. Smolen, J.S., et al., *EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs.* Ann Rheum Dis, 2010. **69**(6): p. 964-75.

6. Knevel, R., et al., *Current evidence for a strategic approach to the management of rheumatoid arthritis with disease-modifying antirheumatic drugs: a systematic literature review informing the EULAR recommendations for the management of rheumatoid arthritis.* Ann Rheum Dis, 2010. **69**(6): p. 987-94.

7. Goekoop-Ruiterman, Y.P., 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): p. 3381-90.

8. Neogi, T., et al., *The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: Phase 2 methodological report.* Arthritis & Rheumatism, 2010. **62**(9): p. 2582-2591.

9. Smolen, J.S. and D. Aletaha, *Monitoring rheumatoid arthritis.* Curr Opin Rheumatol, 2011. **23**(3): p. 252-8.

10. Schoels, M., et al., *Evidence for treating rheumatoid arthritis to target: results of a systematic literature search.* Ann Rheum Dis, 2010. **69**(4): p. 638-43.

11. Schipper, L.G., et al., *Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome.* Rheumatology (Oxford), 2010. **49**(11): p. 2154-64.

12. Molenaar, E.T., et al., *Progression of radiologic damage in patients with rheumatoid arthritis in clinical remission.* Arthritis Rheum, 2004. **50**(1): p. 36-42.

13. Smolen, J.S., et al., *Evidence of radiographic benefit of treatment with infliximab plus methotrexate in rheumatoid arthritis patients who had no clinical improvement: a detailed subanalysis of data from the anti-tumor necrosis factor trial in rheumatoid arthritis with concomitant therapy study.* Arthritis Rheum, 2005. **52**(4): p. 1020-30.

14. Bakker, M.F., 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. **70 (6)**: p. 1099-1103.

15. Berglin, E., et al., *Radiological outcome in rheumatoid arthritis is predicted by presence of antibodies against cyclic citrullinated peptide before and at disease onset, and by IgA-RF at disease onset.* Annals of the Rheumatic Diseases, 2006. **65**(4): p. 453-8.

16. Berglin, E., et al., *Predictors of radiological progression and changes in hand bone density in early rheumatoid arthritis.* Rheumatology, 2003. **42**(2): p. 268-75.

17. Mottonen, T., et al., *Only high disease activity and positive rheumatoid factor indicate poor prognosis in patients with early rheumatoid arthritis treated with "sawtooth" strategy.* Annals of the Rheumatic Diseases, 1998. **57**(9): p. 533-9.

18. Westhoff, G., R. Rau, and A. Zink, *Rheumatoid arthritis patients who smoke have a higher need for DMARDs and feel worse, but they do not have more joint damage than non-smokers of the same serological group.* Rheumatology, 2008. **47**(6): p. 849-54.

19. Van Tuyl, L.H.D., et al., *Evidence for predictive validity of remission on long-term outcome in rheumatoid arthritis: A systematic review.* Arthritis Care and Research, 2010. **62 (1)**: p. 108-117.

20. Corbett, M., et al., *Factors predicting death, survival and functional outcome in a prospective study of early rheumatoid disease over fifteen years.* Br J Rheumatol, 1993. **32**(8): p. 717-23.

21. Welsing, P.M., et al., *The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis.* Arthritis Rheum, 2001. **44**(9): p. 2009-17.

22. Bansback, N., et al., *A prognostic model for functional outcome in early rheumatoid arthritis.* Journal of Rheumatology, 2006. **33**(8): p. 1503-10.

23. Combe, B., et al., *Predictive factors of 5-year health assessment questionnaire disability in early rheumatoid arthritis.* Journal of Rheumatology, 2003. **30**(11): p. 2344-9.

24. Dirven, L., et al., *Towards personalized treatment: Predictors of short-term HAQ response in recent-onset active rheumatoid arthritis are different from predictors of rapid radiological progression.* Scandinavian Journal of Rheumatology, 2012. **41 (1)**: p. 15-19.

25. Lindqvist, E., et al., *Ten year outcome in a cohort of patients with early rheumatoid arthritis: Health status, disease process, and damage.* Annals of the Rheumatic Diseases, 2002. **61 (12)**: p. 1055-1059.

26. van der Kooi, E., et al., *A decrease in disease activity score (DAS) level is associated with a decrease in health assessment questionnaire (HAQ) score, independent of follow-up duration, during 5 years of tightly controlled treatment: results from the BeSt study.* Annals of the Rheumatic Diseases, 2011. **70**(1): p. 168-71.

27. Verstappen, S.M.M., et al., *Functional Health Assessment Questionnaire (HAQ) and psychological HAQ are associated with and predicted by different factors in rheumatoid arthritis.* Journal of Rheumatology, 2007. **34 (9)**: p. 1837-1840.

28. Boyesen, P., et al., *MRI in early rheumatoid arthritis: synovitis and bone marrow oedema are independent predictors of subsequent radiographic progression.* Annals of the Rheumatic Diseases, 2011. **70**(3): p. 428-33.

29. Combe, B., et al., *Prognostic factors for radiographic damage in early rheumatoid arthritis: a multiparameter prospective study.* Arthritis & Rheumatism, 2001. **44**(8): p. 1736-43.

30. Courvoisier, N., et al., *Prognostic factors of 10-year radiographic outcome in early rheumatoid arthritis: a prospective study.* Arthritis Research & Therapy, 2008. **10**(5): p. R106.

31. De Vries-Bouwstra, J., et al., *Using predicted disease outcome to provide differentiated treatment of early rheumatoid arthritis.* Journal of Rheumatology, 2006. **33 (9)**: p. 1747-1753.

32. Dixey, J., C. Solymossy, and A. Young, *Is it possible to predict radiological damage in early rheumatoid arthritis (RA)? A report on the occurrence, progression, and prognostic factors of radiological erosions over the first 3 years in 866 patients from the Early RA Study (ERAS).* The Journal of rheumatology, 2004. **Supplement. 69**: p. 48-54.

33. Fex, E., et al., *Development of radiographic damage during the first 5-6 yr of rheumatoid arthritis. A prospective follow-up study of a Swedish cohort.* British Journal of Rheumatology, 1996. **35 (11)**: p. 1106-1115.

34. Forslind, K., et al., *Prediction of radiological outcome in early rheumatoid arthritis in clinical practice: Role of antibodies to citrullinated peptides (anti-CCP).* Annals of the Rheumatic Diseases, 2004. **63 (9)**: p. 1090-1095.

35. Goronzy, J.J., et al., *Prognostic markers of radiographic progression in early rheumatoid arthritis.* Arthritis & Rheumatism, 2004. **50**(1): p. 43-54.

36. Hetland, M.L., et al., *MRI bone oedema is the strongest predictor of subsequent radiographic progression in early rheumatoid arthritis. Results from a 2-year*

randomised controlled trial (CIMESTRA). Annals of the Rheumatic Diseases, 2009. **68**(3): p. 384-90.

37. Hetland, M.L., et al., *Radiographic progression and remission rates in early rheumatoid arthritis - MRI bone oedema and anti-CCP predicted radiographic progression in the 5-year extension of the double-blind randomised CIMESTRA trial.* Annals of the Rheumatic Diseases, 2010. **69**(10): p. 1789-95.

38. Ichikawa, Y., et al., *Clinical activity after 12 weeks of treatment with nonbiologics in early rheumatoid arthritis may predict articular destruction 2 years later.* Journal of Rheumatology, 2010. **37**(4): p. 723-9.

39. Kaltenhauser, S., et al., *Immunogenetic markers and seropositivity predict radiological progression in early rheumatoid arthritis independent of disease activity.* Journal of Rheumatology, 2001. **28**(4): p. 735-44.

40. Korpela, 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 and Rheumatism, 2004. **50 (7)**: p. 2072-2081.

41. Landewe, R.B.M., et al., *COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention.* Arthritis & Rheumatism, 2002. **46**(2): p. 347-56.

42. Lindqvist, E., et al., *Course of radiographic damage over 10 years in a cohort with early rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2003. **62**(7): p. 611-6.

43. Machold, K.P., et al., *Very recent onset rheumatoid arthritis: clinical and serological patient characteristics associated with radiographic progression over the first years of disease.* Rheumatology, 2007. **46**(2): p. 342-9.

44. Manfredsdottir, V.F., et al., *The effects of tobacco smoking and rheumatoid factor seropositivity on disease activity and joint damage in early rheumatoid arthritis.* Rheumatology, 2006. **45**(6): p. 734-40.

45. Mouterde, G., et al., *Predictors of radiographic progression in the ESPOIR cohort: the season of first symptoms may influence the short-term outcome in early arthritis.* Annals of the Rheumatic Diseases, 2011. **70**(7): p. 1251-6.

46. Nyhall-Wahlin, B.M., et al., *The presence of rheumatoid nodules at early rheumatoid arthritis diagnosis is a sign of extra-articular disease and predicts radiographic progression of joint destruction over 5 years.* Scandinavian Journal of Rheumatology, 2011. **40**(2): p. 81-7.

47. Salaffi, F., et al., *Relationship between time-integrated disease activity estimated by DAS28-CRP and radiographic progression of anatomical damage in patients with early rheumatoid arthritis.* BMC Musculoskeletal Disorders, 2011. **12**: p. 120.

48. Sanmarti, R., et al., *Radiological progression in early rheumatoid arthritis after DMARDS: a one-year follow-up study in a clinical setting.* Rheumatology, 2003. **42**(9): p. 1044-9.

49. Sanmarti, R., et al., *Prognostic factors of radiographic progression in early rheumatoid arthritis: a two year prospective study after a structured therapeutic strategy using DMARDs and very low doses of glucocorticoids.* Clinical Rheumatology, 2007. **26**(7): p. 1111-8.

50. Tanaka, N., et al., *Synovial membrane enhancement and bone erosion by magnetic resonance imaging for prediction of radiologic progression in patients with early rheumatoid arthritis.* Rheumatology International, 2005. **25**(2): p. 103-7.

51. Tengstrand, B., M. Ahlmen, and I. Hafstrom, *The Influence of Sex on Rheumatoid Arthritis: A Prospective Study of Onset and Outcome after 2 Years.* Journal of Rheumatology, 2004. **31 (2)**: p. 214-222.

52. van der Heijde, D.M., et al., *Prognostic factors for radiographic damage and physical disability in early rheumatoid arthritis. A prospective follow-up study of 147 patients.* British Journal of Rheumatology, 1992. **31**(8): p. 519-25.

53. Welsing, P.M.J., et al., *The relationship between disease activity and radiologic progression in patients with rheumatoid arthritis: A longitudinal analysis.* Arthritis and Rheumatism, 2004. **50 (7)**: p. 2082-2093.

54. Kroot, E.J., et al., *The prognostic value of anti-cyclic citrullinated peptide antibody in patients with recent-onset rheumatoid arthritis.* Arthritis Rheum, 2000. **43**(8): p. 1831-5.

55. Matschke, V., et al., *Adverse changes in tendon-muscle physiology and physical function caused by an isolated acute rheumatoid knee effusion: a case study.* Arthritis care & research, 2012. **64**(1): p. 117-21.

56. Plant, M.J., et al., *What factors influence functional ability in patients with rheumatoid arthritis. Do they alter over time?* Rheumatology (Oxford), 2005. **44**(9): p. 1181-5.

57. Lillegraven, S., et al., *What is the clinical relevance of erosions and joint space narrowing in RA?* Nature Reviews Rheumatology, 2012.

58. Landewe, R. and D. van der Heijde, *Joint space narrowing, cartilage and physical function: are we deceived by measurements and distributions?* Postgrad Med J, 2011. **87**(1033): p. 737-8.

59. van der Heijde, D.M., et al., *Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis.* Ann Rheum Dis, 1992. **51**(2): p. 177-81.

60. van der Heijde, D.M., et al., *Disease activity score.* Ann Rheum Dis, 1992. **51**(1): p. 140.

61. Contreras-Yanez, I., et al., *Radiographic outcome in Hispanic early rheumatoid arthritis patients treated with conventional disease modifying anti-rheumatic drugs.* European Journal of Radiology, 2011. **79 (2)**: p. e52-e57.

62. Grigor, C., et al., *Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): A single-blind randomised controlled trial.* Lancet, 2004. **364 (9430)**: p. 263-269.

63. van Vollenhoven, R.F., 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.* Lancet, 2009. **374**(9688): p. 459-66.

64. Masi, A.T., *Articular patterns in the early course of rheumatoid arthritis.* Am J Med, 1983. **75**(6A): p. 16-26.

65. Drossaers-Bakker, K.W., et al., *Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time.* Arthritis Rheum, 1999. **42**(9): p. 1854-60.

66. Goekoop-Ruiterman, Y.P., et al., *Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial.* Ann Intern Med, 2007. **146**(6): p. 406-15.

67. Verstappen S M M, J.J.W.G., van der Veen M J , Heurkens A H M, et al, *Intensive treatment with methotrexate in early rheumatoid rthritis: aiming for remission. Computer AssistedManagement in Early Rheumatoid Arthritis (CAMERA, anopen-label strategy trial).* Ann Rheum Dis 2007. **66**: p. 1443-1449.

68. J Fransen, H.B.M., I Speyer, P L C M van Riel, *Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial.* Ann Rheum Dis, 2005. **64**: p. 1294–1298.

69. Kuriya, B., et al., *Remission in early rheumatoid arthritis -- a comparison of new ACR/EULAR remission criteria to established criteria.* J Rheumatol, 2012. **39**(6): p. 1155-8.

70. Stucki, G., et al., *A self-administered rheumatoid arthritis disease activity index (RADAI) for epidemiologic research. Psychometric properties and correlation with parameters of disease activity.* Arthritis Rheum, 1995. **38**(6): p. 795-8.

71. Bruce, B. and J.F. Fries, *The Stanford Health Assessment Questionnaire: a review of its history, issues, progress, and documentation.* J Rheumatol, 2003. **30**(1): p. 167-78.

72. Bruce, B. and J.F. Fries, *The Stanford Health Assessment Questionnaire: dimensions and practical applications.* Health Qual Life Outcomes, 2003. **1**: p. 20.

73. Redelmeier, D.A. and K. Lorig, *Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology.* Arch Intern Med, 1993. **153**(11): p. 1337-42.

74. Wells, G.A., et al., *Minimum important difference between patients with rheumatoid arthritis: the patient's perspective.* J Rheumatol, 1993. **20**(3): p. 557-60.

75. Lindqvist, E., et al., *Ten year outcome in a cohort of patients with early rheumatoid arthritis: health status, disease process, and damage.* Ann Rheum Dis, 2002. **61**(12): p. 1055-9.

76. Leigh, J.P. and J.F. Fries, *Predictors of disability in a longitudinal sample of patients with rheumatoid arthritis.* Ann Rheum Dis, 1992. **51**(5): p. 581-7.

77. Sherrer, Y.S., et al., *Disability in rheumatoid arthritis: comparison of prognostic factors across three populations.* J Rheumatol, 1987. **14**(4): p. 705-9.

78. Strating, M.M., W.H. Van Schuur, and T.P. Suurmeijer, *Predictors of functional disability in rheumatoid arthritis: results from a 13-year prospective study.* Disabil Rehabil, 2007. **29**(10): p. 805-15.

79. Bombardier, C., et al., *The relationship between joint damage and functional disability in rheumatoid arthritis: a systematic review.* Ann Rheum Dis, 2011.

80. Bombardier, C., et al., *The relationship between joint damage and functional disability in rheumatoid arthritis: a systematic review.* Annals of the Rheumatic Diseases, 2011.

81. Sokka, T., *Radiographic scoring in rheumatoid arthritis: a short introduction to the methods.* Bull NYU Hosp Jt Dis, 2008. **66**(2): p. 166-8.

82. Sharp, J.T., et al., *Methods of scoring the progression of radiologic changes in rheumatoid arthritis. Correlation of radiologic, clinical and laboratory abnormalities.* Arthritis Rheum, 1971. **14**(6): p. 706-20.

83. Sharp, J.T., M.D. Lidsky, and J. Duffy, *Clinical responses during gold therapy for rheumatoid arthritis. Changes in synovitis, radiologically detectable erosive lesions, serum proteins, and serologic abnormalities.* Arthritis Rheum, 1982. **25**(5): p. 540-9.

84. Sharp, J.T., *Radiographic evaluation of the course of articular disease.* Clin Rheum Dis, 1983. **9**(3): p. 541-57.

85. Larsen, A., K. Dale, and M. Eek, *Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films.* Acta Radiol Diagn (Stockh), 1977. **18**(4): p. 481-91.

86. Ravindran, V. and S. Rachapalli, *An overview of commonly used radiographic scoring methods in rheumatoid arthritis clinical trials.* Clin Rheumatol, 2011. **30**(1): p. 1-6.

87. van der Heijde, D.M., et al., *Effects of hydroxychloroquine and sulphasalazine on progression of joint damage in rheumatoid arthritis.* Lancet, 1989. **1**(8646): p. 1036-8.

88. van der Heijde, D.M., et al., *Biannual radiographic assessments of hands and feet in a three-year prospective followup of patients with early rheumatoid arthritis.* Arthritis Rheum, 1992. **35**(1): p. 26-34.

89. Bruynesteyn, K., et al., *Determination of the minimal clinically important difference in rheumatoid arthritis joint damage of the Sharp/van der Heijde and Larsen/Scott scoring methods by clinical experts and comparison with the smallest detectable difference.* Arthritis Rheum, 2002. **46**(4): p. 913-20.

90. Lindqvist, E., et al., *Course of radiographic damage over 10 years in a cohort with early rheumatoid arthritis.* Ann Rheum Dis, 2003. **62**(7): p. 611-6.

91. Guillemin, F., et al., *Prognostic factors for joint destruction in rheumatoid arthritis: a prospective longitudinal study of 318 patients.* J Rheumatol, 2003. **30**(12): p. 2585-9.

92. Wolfe, F. and J.T. Sharp, *Radiographic outcome of recent-onset rheumatoid arthritis: a 19-year study of radiographic progression.* Arthritis Rheum, 1998. **41**(9): p. 1571-82.

93. Uhlig, T., et al., *The course of rheumatoid arthritis and predictors of psychological, physical and radiographic outcome after 5 years of follow-up.* Rheumatology (Oxford), 2000. **39**(7): p. 732-41.

94. Syversen, S.W., et al., *High anti-cyclic citrullinated peptide levels and an algorithm of four variables predict radiographic progression in patients with rheumatoid arthritis: results from a 10-year longitudinal study.* Ann Rheum Dis, 2008. **67**(2): p. 212-7.

95. Forslind, K., et al., *Prediction of radiological outcome in early rheumatoid arthritis in clinical practice: role of antibodies to citrullinated peptides (anti-CCP).* Ann Rheum Dis, 2004. **63**(9): p. 1090-5.

96. Smolen, J.S., et al., *A simplified disease activity index for rheumatoid arthritis for use in clinical practice.* Rheumatology (Oxford), 2003. **42**(2): p. 244-57.

97. Aletaha, D. and J. Smolen, *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. **23**(5 Suppl 39): p. S100-8.

98. Wayman, *Multiple Imputation For Missing Data: What Is It And How Can I Use It?* Paper presented at the 2003 Annual Meeting of the American Educational ResearchAssociation, Chicago, IL., 2003.

99. Schafer and Olsen, *Multiple Imputation for multivariate analysis: Missing data problems. A data analyst's perspective.* Multivariate Behavioral Research, 1998. **33**(4): p. 545-571.

100. Little, R.J.A. and D.B. Rubin, *Statistical analysis with missing data.* . New York, 1987. **John Wiley & Sons**.

101. Barnard and Xiao-Li, *Applications of multiple imputation in medical studies: from AIDS to NHANES.* Statistical Methods in Medical Research, 1999. **8**: p. 17±36.

102. Yuan, Y.C., *Multiple Imputation for Missing Data: Concepts and NewDevelopment (Version 9.0)*

103. Rubin, D., *Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons, Inc.* 1987.

104. Aletaha, D., J. Smolen, and M.M. Ward, *Measuring function in rheumatoid arthritis: Identifying reversible and irreversible components.* Arthritis Rheum, 2006. **54**(9): p. 2784-92.

105. Smolen, J.S., et al., *Estimation of a numerical value for joint damage-related physical disability in rheumatoid arthritis clinical trials.* Ann Rheum Dis, 2010. **69**(6): p. 1058-64.

106. Odegard, S., et al., *Association of early radiographic damage with impaired physical function in rheumatoid arthritis: a ten-year, longitudinal observational study in 238 patients.* Arthritis Rheum, 2006. **54**(1): p. 68-75.

107. Aletaha, D., J. Funovits, and J.S. Smolen, *Physical disability in rheumatoid arthritis is associated with cartilage damage rather than bone destruction.* Ann Rheum Dis, 2011. **70**(5): p. 733-9.

108. Fransen, J., et al., *Feasibility and validity of the RADAI, a self-administered rheumatoid arthritis disease activity index.* Rheumatology (Oxford), 2000. **39**(3): p. 321-7.

109. Keystone, E.C., J. Smolen, and P. van Riel, *Developing an effective treatment algorithm for rheumatoid arthritis.* Rheumatology (Oxford), 2012. **51 Suppl 5**: p. v48-54.

110. Vastesaeger, N., et al., *A pilot risk model for the prediction of rapid radiographic progression in rheumatoid arthritis.* Rheumatology (Oxford), 2009. **48**(9): p. 1114-21.

111. Visser, K., et al., *A matrix risk model for the prediction of rapid radiographic progression in patients with rheumatoid arthritis receiving different dynamic treatment strategies: post hoc analyses from the BeSt study.* Ann Rheum Dis, 2010. **69**(7): p. 1333-7.

112. Saevarsdottir, S., et al., *Predictors of response to methotrexate in early DMARD naive rheumatoid arthritis: results from the initial open-label phase of the SWEFOT trial.* Ann Rheum Dis, 2011. **70**(3): p. 469-75.

113. Bombardier, C., et al., *A Risk Model for the Prediction of Radiographic Progression: Results from SONORA Study.* Arthritis & Rheumatism, 2009. **61:S276.**

114. Fautrel, B., et al., *Matrix to predict rapid radiographic progression of early rheumatoid arthritis patients from the community treated with methotrexate or leflunomide: results from the ESPOIR cohort.* Arthritis Res Ther, 2012. **14**(6): p. R249.

115. Altman, D.G. and G.H. Lyman, *Methodological challenges in the evaluation of prognostic factors in breast cancer.* Breast Cancer Res Treat, 1998. **52**(1-3): p. 289-303.

116. Hayden, J.A., et al., *Identifying phases of investigation helps planning, appraising, and applying the results of explanatory prognosis studies.* J Clin Epidemiol, 2008. **61**(6): p. 552-60.

117. Amaya-Amaya, J., et al., *Usefulness of patients-reported outcomes in rheumatoid arthritis focus group.* Arthritis, 2012. **2012**: p. 935187.

Appendix 1-A

Appendix 1-A

Medline

Database: Ovid MEDLINE(R) <1946 to February Week 5 2012> Search Strategy:

1. arthritis, rheumatoid/ or caplan syndrome/ or felty's syndrome/ or rheumatoid nodule/ or rheumatoid vasculitis/ (76486)
2. ((absence or minimal or low or no) adj5 activit*).ti,ab. (111444)
3. 1 and 2 (692)
4. Recovery of Function/ or "severity of illness index"/ or sickness impact profile/ or disease progression/ (239847)
5. Health Status Indicators/ or (disabilit* or haq or aims2).ti,ab. or ((sharp or heijde or vanderheijde or larson or genant)adj2 (score or scoring or scores)).ti,ab. or ((radiograph* or radiolog*) adj2 (damage or progress*)).ti,ab. or (physical adj2 function).ti,ab. or disability evaluation/ (122350)
6. prognosis/ or treatment outcome/ or morbidity/ or mortality/ (837410)
7. 1 and 4 and 5 and 6 (588)
8. 3 or 7 (1236)

EMBASE

Database: Embase <1980 to 2012 Week 09> Search Strategy:

1. rheumatoid arthritis/ or felty syndrome/ or rheumatoid nodule/ or pneumoconiosis/ or (caplan* adj2 syndrome*).ti,ab. or rheumatoid vasculitis/ (115504)
2. ((absence or minimal or low or no) adj5 activit*).ti,ab. (122039)
3. 1 and 2 (1104)
4. disease course/ or convalescence/ or deterioration/ or disease duration/ or disease exacerbation/ or recurrent disease/ or relapse/ or remission/ or disease severity/ or general condition deterioration/ or general condition improvement/ or good general condition/ or poor general condition/ or expanded disability status scale/ or functional assessment inventory/ or functional independence measure/ or health assessment questionnaire/ or international prognostic scoring system/ or "nottingham extended activities of daily living

scale"/ or oswestry disability index/ or rivermead mobility index/ or short form 36/ or sickness impact profile/ (773364)

5. health survey/ or (disabilit* or haq or aims2).ti,ab. or ((sharp or heijde or vanderheijde or larson or genant) adj2 (score or scoring or scores)).ti,ab. or ((radiograph* or radiolog*) adj2 (damage or progress*)).ti,ab. or (physical adj2 function).ti,ab. or disability severity/ (244799)
6. prognosis/ or treatment outcome/ or disease free interval/ or outcome assessment/ or outcomes research/ or exp treatment failure/ or morbidity/ or mortality/ (1482175)
7. 1 and 4 and 5 and 6 (909)
8. 3 or 7 (1960)

Appendix 1-B

Appendix 1-B

Characteristics of included studies

No	Study	Year	Country	Design/population	Sample size	Disease duration	Outcome	Measure of DA
1	Bakker	2011	Netherlands	RCT (CAMERA)	299	<1 yr	Damage	DAS28
2	Bansback	2006	UK	ERAS	985	<2 yr	Function	DAS28, SJC, TJC
3	Berglin	2006	Sweden	ERA cohort	138	<1yr	Damage	DAS28, SJC, TJC
4	Berglin	2003	Sweden	ERA cohort	43	<1yr	Damage	DAS28, SJC, TJC
5	Boyensen	2011	Norway	ERA cohort	84	<1yr	Damage	DAS28, SJC, TJC
6	Combe	2003	France	ERA cohort	191	<1yr	Function	SJC, TJC, RAI, DAS
7	Combe	2001	France	ERA cohort	191	<1yr	Damage	SJC, TJC, RAI, DAS
8	Courvoisier	2008	France	ERA cohort	191	<1yr	Damage	DAS, TJC, SJC
9	Corbett	1993	UK	ERA cohort	102	<1yr	Function	AJC
10	De Vries-Bouwstra	2006	Netherlands	Leiden ERA	152	<2yr	Damage	SJC
11	Dirven	2012	Netherlands	RCT (BeST)	497	<2yr	Function	RAI
12	Dixey	2004	UK	ERAS	866	<2yr	Damage	Joint Score
13	Fex	1996	Sweden	ERA cohort	113	<2yr	Damage	AJC, RAI
14	Forslid	2004	Sweden	BARFOT	379	<1yr	Damage	DAS28
15	Goronzy	2004	USA	ERA cohort	111	<1yr	Damage	SJC
16	Hetland	2010	Denmark	CIMESTRA	130	<6mo	Damage	DAS28
17	Hetland	2009	Denmark	CIMESTRA	160	<6mo	Damage	DAS28
18	Ichikawa	2009	Japan	ERA cohort	55	<2yr	Damage	DAS28, SJC
19	Kaltenhauser	2001	Germany	ERA cohort	87	<2yr	Damage	SJC
20	Korpela	2004	Finland	FIN-RACo	195	<2yr	Damage	SJC,TJC
21	Kroot	2000	Netherlands	ERA cohort	237	<1yr	Damage, Function	DAS
22	Landewe	2002	Netherlands	COBRA	115	<2yr	Damage	DAS28
23	Lindqvist	2003	Sweden	ERA cohort	183	<2yr	Damage	AJC
24	Lindqvist	2002	Sweden	ERA cohort	183	<2yr	Function	AJC
25	Machold	2007	Austria	VERA cohort	138	<3mo	Damage	SJC, TJC, DAS28
26	Manfredsdottir	2006	Iceland	ERA cohort	100	<1yr	Damage	SJC
27	Mottonen	1998	Finland	ERA	142	<2yr	Damage	SJC, TJC, Mallya
28	Mouterde	2011	France	ESPOIR	736	<6mo	Damage	DAS28
29	Nyhall-Wahlin	2011	Sweden	BARFOT cohort	336	<1yr	Damage	DAS28
30	Salaffi	2011	Italy	ERA cohort	59	<1yr	Damage	DAS28

No	Study	Year	Country	Design/population	Sample size	Disease duration	Outcome	Measure of DA
31	Sanmarti	2007	Spain	ERA cohort	105	<2yr	Damage	DAS28
32	Sanmarti	2003	Spain	ERA cohort	60	<2yr	Damage	DAS28
33	Tanaka	2005	Japan	ERA cohort	130	<1yr	Damage	SJC, TJC
34	Tengstrand	2004	Sweden	BARFOT	844	<1yr	Damage	DAS28
35	van der Heijdeh	1992	Netherlands	ERA	147	<1yr	Damage	DAS
36	van der Kooi	2011	Netherlands	BeST trial	508	<2yr	Function	DAS
37	Verstappen	2007	Netherlands	2 trials*	112	<1yr	Function	Thompson score **
38	Welsing	2004	Netherlands	UMCN, COBRA	185+152	<1yr	Damage	DAS28
39	Welsing	2001	Netherlands	UMCN	203	<1 yr	Function	DAS
40	Westhoff	2008	Germany	ERA cohort	896	<2yr	Damage	DAS28

RAI=Ritchie Articular Index; AJC=Active Joint Count; SJC=Swollen Joint Count; TJC=Tender Joint Count; Dutch Behandel Strategieen (BeSt) study; ERAS= Early RA Study; CAMERA: Computed Management of ERA; * Utrecht Rheumatoid Arthritis Cohort study group and CAMERA; **Thompson Joint Score: weighted score of both tender and swollen joints, total score 0–534; UMCN= University Medical Center Nijmegen; COBRA= the Maastricht Combination Therapy in RA

Appendix 1-C

Prognostic Studies. Outcome: radiographic damage (studies demonstrating significant association between disease activity and damage are highlighted (grey))

No	Study	Year	Country	Design	Sample size	Disease duration	F/U	Outcome	Measure of DA in final model (if applicable)
Outcome: Joint damage score (continuous variable)									
1	Berglin	2006	Sweden	ERA cohort	138	<1yr	2 yr	LS	SJC
2	Kroot	2000	Netherlands	ERA cohort	237	<1yr	6 yr	SHS	DAS
3	van der Heijdeh	1992	Netherlands	ERA	147	<1yr	2 yr	SHS	SJC
4	Manfredsdottir	2006	Iceland	ERA cohort	100	<1yr	2 yr	SHS	SJC
5	Tengstrand	2004	Sweden	BARFOT	844	<1yr	2 yr	LS	NA
Outcome: Joint damage score change (continuous variable)									
1	Bakker	2011	Netherlands	RCT (CAMERA)	299	<1 yr	5 yr	Mean annual ΔSHS	6m EULAR resp
2	Ichikawa	2009	Japan	ERA cohort	55	<2yr	8 yr	ΔSHS	SJC at 12 wks
3	De Vries-Bouwstra	2006	Netherlands	Leiden ERA	152	<2yr	1 yr	ΔSHS	SJC
4	Machold	2007	Austria	VERA cohort	138	<3mo	3 yr	ΔLS	Cumulative SJC
5	Welsing	2004	Netherlands	UMCN, COBRA	185+152	<1yr	9, 6 y	ΔSHS	DAS28
6	Berglin	2003	Sweden	ERA cohort	43	<1yr	2 y	ΔLS	6 mo EULAR resp, SJC
7	Landewe	2002	Netherlands	COBRA	115	<2yr	4-7 y	Annual ΔSHS	DAS28
8	Kroot	2000	Netherlands	ERA cohort	237	<1yr	3 yr	ΔSHS	DAS
9	van der Heijdeh	1992	Netherlands	ERA	147	<1yr	2 yr	ΔSHS	Mean DAS (3-6mo)
10	Nyhall-Wahlin	2011	Sweden	BARFOT cohort	336	<1yr	5 yr	ΔSHS	NA
11	Boyensen	2011	Norway	ERA cohort	84	<1yr	3 yr	ΔSHS	NA
12	Hetland	2010	Denmark	CIMESTRA	130	<6mo	5 yr	ΔSHS	DAS28
13	Hetland	2009	Denmark	CIMESTRA	160	<6mo	2 yr	ΔSHS	DAS28
14	Kaltenhauser	2001	Germany	ERA cohort	87	<2yr	2 yr	Annual ΔSHS	SJC
Outcome: Joint damage progression (dichotomous variable)									
1	Salaffi	2011	Italy	ERA cohort	59	<1yr	3 yr	ΔSHS>9.5 (SDD)	DAS28
2	Westhoff	2008	Germany	ERA cohort	896	<2yr	3 yr	R score ≥ 7 (SDD)	DAS28
3	Berglin	2006	Sweden	ERA cohort	138	<1yr	2 yr	ΔLS > medianΔ ; LS	6m Tx resp, SJC
4	Mottonen	1998	Finland	ERA	142	<2yr	6 yr	LS>50*	Mallya
5	Mouterde	2011	France	ESPOIR cohort	736	<6mo	6 mo	ΔSHS >1 (SDD)	NA
6	Hetland	2010	Denmark	CIMESTRA	130	<6mo	5 yr	SHS>0	NA
7	Courvoisier	2008	France	ERA cohort	191	<1yr	10 yr	ΔSHS ≥5 (OMERACT MCID)	NA
8	Sanmarti	2007	Spain	ERA cohort	105	<2yr	2 yr	ΔLS>4 (MCID)	NA

No	Study	Year	Country	Design	Sample size	Disease duration	F/U	Outcome	Measure of DA in final model (if applicable)
9	Tanaka	2005	Japan	ERA cohort	130	<1yr	10 yr	Severe ΔSHS ($^{\delta}$ damage)	NA
10	Dixey	2004	UK	ERAS	866	<2yr	3 yr	Severe erosion (NR)	SJC
11	Forslind	2004	Sweden	BARFOT	379	<1yr	2 yr	SHS>10; ΔSHS>8(both median)	NA
12	Goronzy	2004	USA	ERA cohort	111	<1yr	2 yr	Erosion	NA
13	Korpela	2004	Finland	FIN-RACo (RCT)	195	<2yr	5 yr	x-ray progression**	SJC,TJC
14	Lindqvist	2003	Sweden	ERA cohort	183	<2yr	10 yr	ΔLS ≥11 units (SDD)	NA
15	Sanmarti	2003	Spain	ERA cohort	60	<2yr	1 yr	ΔLS>2 (MCID)	NA
16	Combe	2001	France	ERA cohort	191	<1yr	3 yr	SHS> 4(median), ΔSHS> 3.4 (>95%CIof ΔSHS)	NA
17	Fex	1996	Sweden	ERA cohort	113	<2yr	5 yr	ΔSHS (highest 1/3 c/t the rest)	NA

NR=not reported, c/t= compare to , NA=not applicable, *34% had progressed to Larsen >50 this was chosen as the limit; **progression was determined based on Larsen system, joint score was categorized into four groups; SDD= smallest detectable difference, MCID=minimal clinically important difference; LS=Larsen Score, Resp=response

Appendix 1-D

Appendix 1-D

Outcome: joint damage score at end point (continuous variable)

No.	Study year	Analysis	Outcome	Predictors	β (95% CI)	p-value	Comments
1	**Berglin 2006**	Linear regression	LS at 2y	LS	β range 0.95-0.99	all<.0001	They reported 4 models and have included one autoantibody in each
				SJC	β range 0.20-0.31	all<.05	
				RF-IgA	4.0	<0.01	
				RF-IgG	2.8	<0.05	
				RF-IgM	4.2	<0.05	
				anti-CCP	2.8	<0.05	
2	**Manfredsdottir 2006**	Linear regression	SHS at 2y	Age	0.02(-0.09,0.12)	0.76	
				Gender	1.72 (-1.7,5.14)	0.32	
				SJC at 6 mo	0.09 (-0.2,0.38)	0.56	
				IgA RF	5.91 (2.46,9.36)	0.001	
3	**Tengstrand 2004**	Linear regression	LS at 2y	Smoking		NS	In men
				RF+		<.01	
				HAQ		NS	
		Linear regression		Dis duration		0.05	In women
				RF+		<.01	
				CRP		<.01	
4	**Kroot 2000**	Linear regression	SHS at 3y	Male†	-0.243(0.392)		Adjusted R^2=0.46
				Age	0.002 (0.012)		
				RF +	1.964 (0.461)‡		
				DAS	0.389 (0.148)‡		
				HLA-DR4	0.562 (0.358)		
				Anti-CCP +	0.209 (0.369)		
				SHS BSL	0.932 (0.082)‡		
		Linear regression	SHS at 6y	Male†	-0.049 (0.464)		Adjusted R^2=0.35
				Age	0.013 (0.016)		
				RF +	2.477 (0.596)‡		
				DAS	0.370 (0.199)		
				HLA-DR4	0.289 (0.463)		
				Anti-CCP +	0.918 (o.477)‡		

No.	Study year	Analysis	Outcome	Predictors	β (95% CI)	p-value	Comments
				SHS BSL	0.900(0.105)‡		
5	**Van de Hejde**		SHS at 2y	RF	positive		R^2=0.31
	1992			ESR	positive		
				SJC	positive		
				DR2	negative		

*β coefficient; † β coefficient (standard error); SJC=Swollen Joint Count; RF=Rheumatoid Factor; DAS=Disease Activity Score; NS=not significant; SHS: Sharp van der Heijde; LS: Larsen Score; BSL: baseline; ‡ p<0.0

Appendix 1-D (continued)

Outcome: x-ray score progression (Δscore) at end-point (continuous variable):

No	Study year	Analysis	Outcome	Predictors	β (95% CI)	p-value	comments
1	**Bakker 2011**	Linear regression	Mean annual SHS progression over 5y	Age	-0.001(-0.01, 0.01)	0.77	Used multivariate imputation analysis. Results were similar after imputation.
				Gender	-0.12 (-0.37, 0.12)	0.33	
				Tx strategy	0.17 (-0.06, 0.40)	0.15	
				RF +	0.35 (0.11, 0.60)	0.01	
				Damage	0.07 (0.04, 0.10)	0.000	
				6m EULAR good resp [c]	-0.43 (-0.74,-0.11)	0.01	
				6m EULAR mod resp [c]	-0.09 (-0.33, 0.21)	0.55	
2	**Nyhall-Wahlen 2011**	Linear regression	x-ray progression at 5y	RF+	5.73 (1.58, 13.04)	0.12	
				Anti-CCP +	11.57 (4.50, 18.64)	0.001	
				SHS	0.69 (0.22, 1.17)	0.004	
				R nodule	0.50 (7.20, 8.21)	0.90	
3	**Boyesen 2011**	Linear regression	x-ray progression at 3y	Baseline predictors	Model 1* Model 2*	Model 3*	When used time integrated (1 year-AUC)values for the same variables, both MRI bone marrow edema and tenosynovitis were significant same as anti-CCP and ESR but R^2 was lower compare to shown model 1-3
				MRI BM edema	0.17(0.28)	0.26(0.06)	
				MRI tenosynovitis	0.19(0.19) 0.34(0.02)		
				Anti-CCP+	0.28(0.04)		
				Female	-0.13(0.32)	-0.17(0.21)	
				Age	-0.07(0.62)	-0.02(0.87)	
4	**Hetland 2010**	Linear Mixed model	ΔSHS at 5 y	Gender (male)	-0.13 (-3.09, 2.84)	0.93	
				Age	0.08 (-0.04, 0.19)	0.2	
				DAS28	-0.07 (-1.20, 1.05)	0.9	
				MRI erosion score	0.13 (-0.82, 1.07)	0.8	
				MRI synovitis score	-0.29 (-0.91, 0.33)	0.4	
				MRI BM edema sc	0.83 (0.45, 1.22)	<.001	
				Anti-CCP+	3.00 (0.33, 5.70)	0.03	
				SHS	0.24 (0.03, 0.45)	0.02	
5	**Ichikawa 2010**	Linear regression	ΔSHS at 8 y	TSS BSL	0.59, 0.33ψ	0.003	
				SJC at 12wks	1.36, 0.28 ψ	0.02	
				CRP 12 wks	1.76, 0.23 ψ	0.05	
				Pain 12 wks	0.31, 0.28 ψ	0.02	
6	**Hetland 2009**	Linear regression	ΔSHS at 2 y	Bone marrow edema	0.75 (0.55- 0.94)	<0.001	DAS28, TJC, SJC NS in univariate But DAS28 incld in MV and was NS

No	Study year	Analysis	Outcome	Predictors	β (95% CI)	p-value	comments
7	**de VRIES-Bouwstra 2006**	Linear regression	x-ray progression at 1 y	Age	-0.1	0.39	
				Sex	1.2	0.71	
				VAS disease activity	0.9	0.06	
				AMS duration	0.0	0.98	
				SJC	1.7	0.01	
				Ritchie score	-0.0	0.89	
				MTP comp pain	6.0	0.09	
				RF+	10.2	0.003	
				ESR	0.1	0.28	
				Shared epitope	0.7	0.83	
				HAQ	-3.4	0.21	
				SHS	0.3	0.31	
				Presence erosion	1.2	0.97	
8	**Machold 2004**	Linear regression stepwise	x-ray progression over 3 y	RF+	0.321*	0.05	Model 1 (Adjusted R^2=0.32)
				Anti-CCP+	0.314*	0.05	
				Time in LDA	-0.39*	<.0001	Model 2 (Adjusted R^2=0.61)
				Cumulative SJC	0.26*	0.01	
				Cumulative CRP	0.19*	0.05	
9	**Welsing 2004**	GEE	x-ray progression	Intercept	4.2 (-14.2, 22.5)	0.66	UMCN Cohort
				Time	-6.7 (-12.1,-1.3)	0.01	
				$Time^2$	0.5 (0.0, 0.9)	0.03	
				RF	32.6 (17.2, 48.0)	<0.0001	
				SHS BSL	0.5 (0.2, 0.8)	0.004	
				SHS BSL-time	-0.1 (-0.1, -0.0)	0.04	
				RF-time	-3.4 (-5.3,-1.4)	0.001	
				Mean DAS	5.4(2.1,8.6)	0.001	
				SD of mean DAS	20.2 (7.2,33.2)	0.002	
				Previous SHS	1.1 (1.0, 1.2)	<0.0001	
			x-ray Progression	Intercept	-0.1 (-6.3, 6.1)	0.98	COBRA
				Time	-2.3 (-6.1, 1.5)	0.21	
				Time2	0.4 (-0.2, 0.8)	0.24	
				Treatment	-1.7 (-5.1, 1.7)	0.33	
				RF +	3.2 (0.4, 6.0)	0.02	
				SHS BSL	-0.1 (-0.2, 0.0)	0.15	
				Treatment-time	0.3 (-1.1, 1.7)	0.63	

No	Study year	Analysis	Outcome	Predictors	β (95% CI)	p-value	comments
				SHS BSL-time	-0.1 (-1.1, 0.9)	0.88	
				RF-time	-0.04 (-0.08,-0.00)	0.06	
				DAS28	1.4 (0.8, 2.0)	<0.0001	
				Previous SHS	1.2 (1.1, 1.3)	<0.0001	
10	**Berglin 2003**	Linear regression	x-ray progression at 2 y	Rx response†: good	-14.64(-24.22, -5.06)	0.004	
				Intermediate	-11.23 (-19.80, -2.65)	0.012	
				SE+	9.28 (1.88, 16.68)	0.02	
				Larsen score	-0.42 (-0.74, -0.10)	0.01	
				SJC	0.68 (2.16 E-02, 1.35)	0.04	
				HAQ	3.07 (-6..82, 12.96)	0.53	
				CRP	0.18 (-1.63E-02, 0.38)	0.07	
11	**Landewe 2002**	GEE	Annual x-ray progression	Tx (SSZ vs COBRA)	-3.2(-5.6. -0.8)	0.010	
				RF+	3.6 (1.2, 5.0)	0.004	
				SHS BSL	0.20 (0.08, 0.32)	0.001	
				DAS28 BSL	1.20 (0.20, 2.20)	0.050	
				Age	0.06 (-0.06, 0.18)	0.288	
				Male	0.82 (-1.90, 3.58)	0.551	
				Disease duration	-0.04 (-0.28, 0.20)	0.720	
				HLA-DR4	-0.45 (-1.35, 0.45)	0.591	
12	**Kaltenhauser 2001**	mixed effect regression with a random	Annual x-ray Progression over 4 y intercept	Intercept	7.23 (4.20, 10.26)	<.001	SJC at BSL and 6m and Ritchie index were NS in multivariate model
				Time	-3.06 (-5.59, -0.52)	0.02	
				SE +	3.24 (0.87, 5.61)	0.01	
				RF level	0.01 (0.00, 0.02)	0.01	
				male	2.92 (5.95, 0.11)	0.06	
13	**Kroot 2000**	Linear regression	ΔSHS at 3 y	Male ϒ	-0.206(0.358)		
				Age	<.001(0.012)		
				RF+	1.928(0.458)‡		
				DAS	0.392(0.148)‡		
				HLA-DR4	0.576(0.357)		
				Anti-CCP+	0.226(0.368)		
			ΔSHS at 6 y	Male ϒ	-0.004(0.462)		
				Age	0.011(0.016)		
				RF+	2.432(0.594)‡		
				DAS	0.376(0.199)		
				HLA-DR4	0.411(0.460)		

No	Study year	Analysis	Outcome	Predictors	β (95% CI)	p-value	comments
				Anti-CCP+	0.933(o.477)‡		
14	**Van der Hejde 1992**	Linear regression	x-ray Progression at 2 y	DAS3-6m (mean) RF + Sharp score BSL HLA-DR2	positive Positive Positive Positive	R^2=0.39	significant variables kept in final model but no parameter estimate/ p-value were reported

*β (p-value); † Rx response= EULAR treatment response at 6 mo, poor response was reference; [Y] Coefficient (standard error); ψ estimates and standardized estimates; NS= not significant; SHS= Sharp van der Hejde score; [€] poor EULAR response was the reference category; ‡p<0.05; VAS: Visualized Analogue Scale; AMS=morning stiffness; MTP= metatarsophalangeal, comp=compression; SE=shared epitope ; GEE=Generalized Estimating Equation; SSZ=Sulfasalazine;

Appendix 1-D (continued)

Outcome: X-ray progression at end point (dichotomous variable)

No.	Study year	Analysis	Outcome	Predictors	OR (95% CI)	p-value	comments
1	**Mouterde 2011**	Logistic regression	x-ray progression at 6 mo	Anti-CCP+	2.94 (1.88 , 4.59)	<0.001	
				TSS BSL	2.56 (1.64 , 3.99)	<0.001	
				HLADRB1*01or 04double	2.67 (1.30 , 5.50)	0.008	
				HLADRB1*01or04 singl	1.86 (1.13, 3.05)	0.01	
				ESR>median	2.04 (1.32, 3.14)	0.001	
				Season (w&s) vs(s&a)	1.66 (1.07, 2.59)	0.02	
2	**Salaffi 2011**	Logistic regression	x-ray progression	time-integration of the DAS28-CRP over 3yr		< .0001	
				Anti-CCP		< .0001	
				IgM-RF		0.0009	
				Joint damage		0.004	
3	**Hetland 2010**	Logistic regression	x-ray progression at 5 years	Bone marrow edema	1.44 (0.95, 2.2)	0.09	
				Anti-CCP	4.03 (1.6, 9.8)	0.002	
				SHS	1.12 (1.0, 1.2)	0.006	
4	**Courvoisier 2008**	Logistic regression	SHS at 10 y	Erosion score BSL	1.73, 5.64 (1.78, 17.86) †	NR	DAS, TJC, SJC, Ritchi index NS in univariate analysis
				ACPA‡	1.35, 3.87 (1.17, 12.75) †	NR	
				ESR	1.17, 3.20 (1.17, 8.78) †	NR	
5	**Westhoff 2008**	Logistic regression	x-ray progression	male gender	1.6 (1.1, 2.3)		
				RF +	1.6 (1.1, 2.4)		
				↑ CRP (>15 vs <5 mg/l)	2.6 (1.2, 4.3)		
				DAS28 (>5.1 vs <3.2)	1.9 (1.1, 3.1)		
				disease duration	1.03 (1.01, 1.05)		
				BMI <25 vs >30	2.6 (1.5, 4.4)		
6	**Sanmarti 2007**	Logistic regression	ΔLarsen score >4 in 2 y	Female	5.48 (1.07-28.17)	0.04	DAS28, TJC, SJC NS in univariate
				DRB1*04+	3.15 (1.10-9.00)	0.03	
				Anti-CCP +	3.63 (0.91-14.46)	0.06	
7	**Berglin 2006**	Logistic regression	x-ray progression at 2 y	Anti-CCP+	5.4 (1.7, 7.0)	<.01	
				SJC	1.1 (1.0, 1.2)	<.05	
				6mo therapeutic responseμ	0.3 (0.14, 0.8)	<.05	
				IgA-RF	9.8 (2.1, 45.5)	<.01	
				ESR	1.0 (1.00, 1.04)	0.05	
8	**Tanaka 2005**	Logistic	severe x-ray	MRI score	3.59 (1.53,8.39)		

No.	Study year	Analysis	Outcome	Predictors	OR (95% CI)	p-value	comments
		regression	progression	CRP	2.86 (1.01,5.88)		
				RF +	2.07 (1.01,3.11)		
9	**Dixey 2004**	Logistic regression	no/ mild/mod vs severe erosion at 3yr	SJC (1-3)[£]	1.7 (0.9,3.2) to 0.6 (0.3, 1.4)		
				nodule	2.6 (1.2, 5.5)		
				Larsen score (1-5) [£]	1.7 (0.7, 3.9) to 2.2 (20.5, 133.2)		
10	**Forslind 2004**	Logistic regression	M1:end point LS of ≥10	Baseline Larsen score	2.7, 14.9 (8-27.6) †	0.0005	DAS28 NS in univariate
				Anti-CCP+	1.5, 4.7 (2.5-8.7) †	0.0005	
				ESR	0.7, 2.0 (1.1-3.5) †	0.02	
			M2:ΔLS from BSL >8	Baseline Larsen score	2.2, 9.3 (5.3-16.1) †	0.0005	
				Anti-CCP+	1.1, 3.0 (1.7-5.2) †	0.0005	
				ESR	0.6, 1.8 (1.0-3.1) †	0.04	
11	**Gorozny 2004**	Logistic regression	erosion at 2 year	RF +	2.9, (1.3, 6.5)	NR	TJC and SJC were NS in univariate
				Erosion	4.0, (1.6, 9.7)	NR	
12	**Korpela 2004**	ordered Logistic regression	x-ray progression at 5 yrs	RF+	2.75 (1.46, 5.17)	0.002	
				Mono vs combo Rx	2.53 (1.44, 4.45)	0.001	
				Disease duration	1.11 (1.04, 1.17)	0.001	
				ESR	1.02 (1.00, 1.03)	0.01	
				TJC	1.02 (0.98,1.06)	0.32	
				Age	0.99 (0.96, 1.02)	0.42	
				SJC	0.99 (0.93, 1.05)	0.78	
				Female	0.74 (0.41, 1.32)	0.31	
13	**Lindqvist 2003**	Logistic regression	M1: x-ray prog 0-5 y	Mean ESR over 3m	0.08, 1.08 (1.04, 1.13) †	NR	
				RF	-1.18,0.31 (0.11, 0.87) †	NR	
				Epitope	-1.93,0.15 (0.03, 0.64) †	NR	
			M2: x-ray prog 5-10 y	Mean ESR over 3m	0.04,1.04 (1.01, 1.07) †	NR	
				RF	-1.47,0.19 (0.09, 0.63) †	NR	
14	**Sanmarti 2003**	Logistic regression	ΔLarsen score >2	disease duration	1.15 (1.03,1.28)		
				VAS pain	1.02 (1.02,1.09)		
				Larsen score	1.06 (1.06,1.55)		
15	**Combe 2001**	Logistic regression	M1: SHS>4	Baseline sharp score	3.4, 31.1 (10.2, 95)†	NR	DAS,TJC, SJC, HAQ were not significant
				RF+	1.1, 2.9 (0.9, 9.2) †	NR	
				HLADRB1*04	1.1, 2.9 (1.0, 8.0) †	NR	
				Pain ≥59	0.9, 2.4 (0.8, 6.6) †	NR	
			M2: x-ray progression	Erosion score≥1	1.6, 5.1 (2.2, 12.1) †	NR	
				RF+	1.3, 3.9 (1.4, 10.6) †	NR	
				ESR	1.2, 3.4 (1.4, 8.5) †	NR	

No.	Study year	Analysis	Outcome	Predictors	OR (95% CI)	p-value	comments
				DRB1*4 +	1.1, 2.9 (1.2, 7.0) †	NR	
16	**Fex 1996**	Logistic regression	x-ray progression at 5 years	1st y Progression rate	0.23Ŧ	0.0005	Active joint count, Ritchie index were included but NS
				Female	-2.59 Ŧ	0.01	
				ESR	0.03 Ŧ	0.03	
17	**Mottonen 1994**	Logistic regression	LS prog > 50 at the latest visit (6 y)	Female	1.7 (0.7, 4.5)		
				Age	1.5 (0.5, 4.1)		
				DR4 +	0.7 (0.3, 1.7)		
				RF +	2.3 (1.0, 5.3)		
				Mayalla score	4.7 (2.0, 11.0)		
				Presence of erosion	0.9 (0.4, 2.1)		
				Dis dur at latest visit	1.0 (0.9, 1.1)		

*HR(95%CI); † Coefficient, OR(95%CI); NR=Not Reported;‡ ACPA: anti-citrullinated protein antibody; £All continuous variables were categorized into quartiles; Ŧ regression coefficient; NS=not significant; µ: no vs mod/good response; NS=not significant; SHS= Sharp van der Hejde score; VAS: Visualized Analogue Scale; SE=shared epitope ; RF=Rheumatoid Factor; DAS=Disease Activity Score; NS=not significant; SHS= Sharp van der Heijde; LS= Larsen Score; BSL= baseline; prog=progression; ‡ p<0.0

Appendices 2-A to L

Appendix 2-A

Comparing patients with available main predictor and outcomes with who missed these variables

Variables	LDAY1 and HAQY3 N=696	No LDAY1 or HAQY3 N=288	LDAY1and x-ray Progression N=654	No LDAY1 or x-ray Progression N=330
Female, n (%)	**519 (75%)**	**189(65%) ‡**	**485(74%)**	**223 (68%) ‡**
Age (yr) *	53.2 (14.0) [53, 43-64]	52.3 (16.5) [51.5, 41.0-65.0]	**52.02 (14.12)** **[51.0, 42-62]**	**54.7 (15.9)‡** **[56.0, 43-67]**
Disease duration (day)*	155.7 (90.9) [138.0, 88-211]	160.0 (105) [136.0, 90.0-218]	155.6 (91.9) [138.0, 8-211]	159.6(101.6) [136.0, 89-215]
RF +ve, n (%)	410 (60%)	183(64%)	387 (60%)	326 (63%)
Anti-CCP +ve, n(%)	303(56%)	118(52%)	270 (53%)	265 (54%)
SDAI *	30.1(16.1) [28.2, 17.8-40.0]	31.4 (17.7) [28.0, 16.0-44.5]	30.0 (16.1) [27.4,17.8-40.8]	31.6 (17.6) [29.8, 16.8-43.9]
SJC *	9.2 (6.9) [8.0, 3.0-14.0]	9.9 (7.4) [9.0,4.0-16.0]	**9.1 (7.0)** **[7.0,3.0-14.0]**	**10.1 (7.2) †** **[10.0, 4.0-15.0]**
TJC *	10.1(7.9) [9.0, 3.0-16.0]	10.2 (8.2) [8.0, 3.0-16.0]	10.0(7.8) [9.0,3.0-16.0]	10.4(8.2) [9.0, 3.0-16.0]
MD global [0-10] *	4.8 (2.03) [5.0, 3.0-6.0]	4.8(2.1) [5.0, 3.0-7.0]	4.8 (2.1) [5.0, 3.0-6.0]	4.8 (2.1) [5.0, 3.0-6.0]
Patient global [0-10] *	4.6(2.4) [5.0, 3.0 -7.0]	4.8(2.4) [5.0, 3.0-6.0]	4.7 (2.4) [5.0, 3.0-6.0]	4.8(2.5) [5.0, 3.0-7.0]
CRP mg/dl *	**1.3 (1.4)** **[0.8, 0.8-1.0]**	**1.6 (1.7)‡** **[0.8, 0.8-1.3]**	**1.3 (1.3)** **[0.8, 0.8-1.0]**	**1.6 (1.8)‡** **(0.8, 0.8-1.25]**
DAS28 *	4.9 (1.3) [4.9, 3.9-5.8]	4.9 (1.4) [5.0, 3.9-6.0]	4.8 (1.3) [4.9, 3.9-5.8]	5.0(1.4) [5.0, 3.9-5.9]
Sharp Score*	4.9 (6.4) [3.0, 0.0-7.0]	5.9 (10.9) [3.0, 1.0-7.0]	5.0 (7.3) [3.0, 0.0-6.0]	5.1(7.2) [3.0, 1.0-7.0]
HAQ *	0.98 (0.7) [0.93, 0.38-1.5]	1.0(0.8) [1.0, 0.4-1.6]	1.0(0.7) [0.9, 0.4-1.5]	1.0(0.8) [1.0, 0.4-1.6]

*mean(sd)[median,95%CI], † P=0.05, ‡p<0.05; LDAY1=LDA at year 1; HAQY3=HAQ at year 3

Appendix 2-B

Table 1: Recruitment and cohort maintenance

	Baseline	Year 1	Year 2
Approached and Recorded	1101		
Consented	984		
Completed		905	846
Deceased		10	15
Withdrawal		20	37
Lost to Follow-up		13	24
Changed Rheumatologist		11	23
Status Unknown		25	39

Appendix 2-C

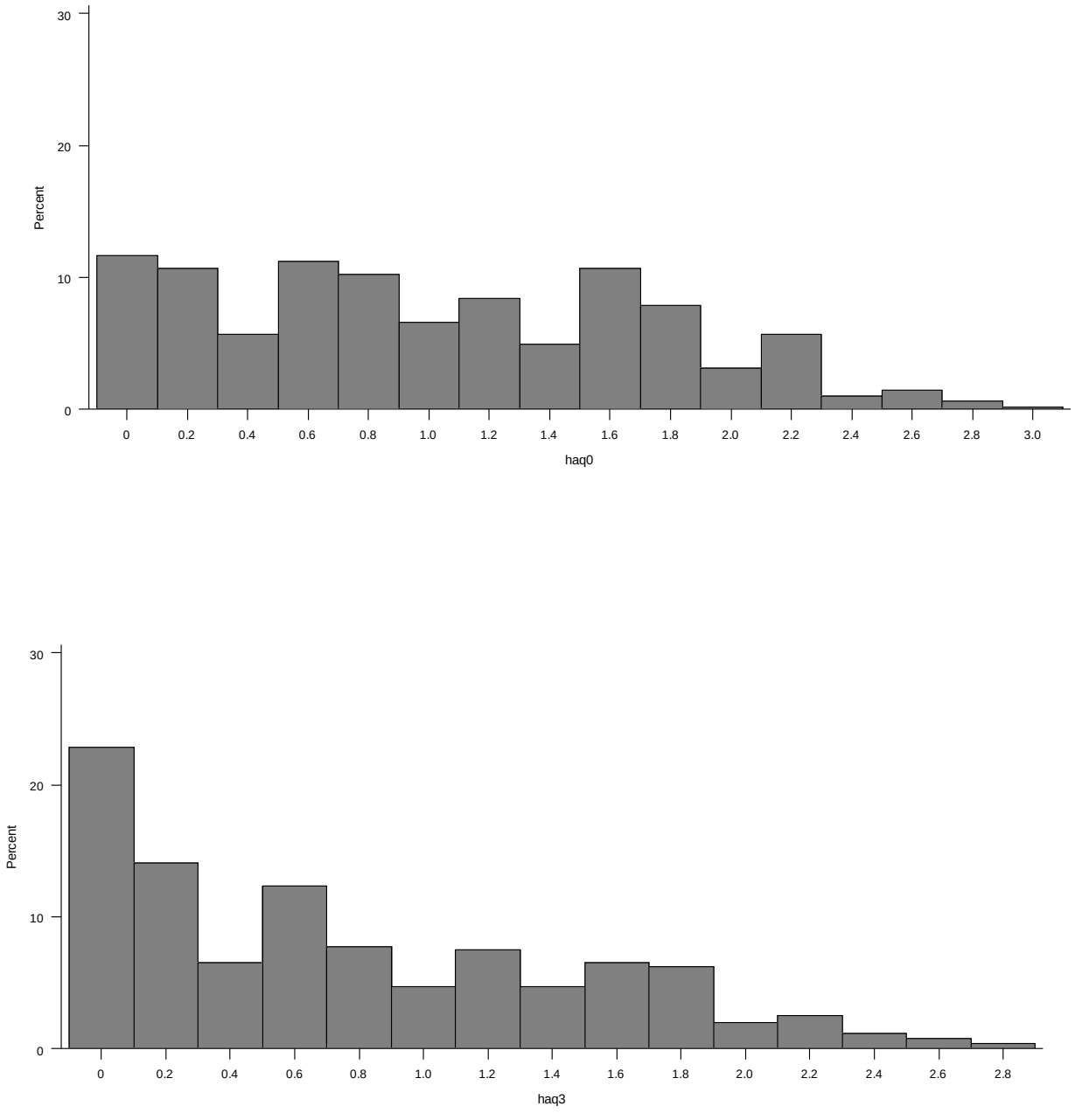

Figure 1- HAQ distribution at baseline (top) and year 3 (bottom)

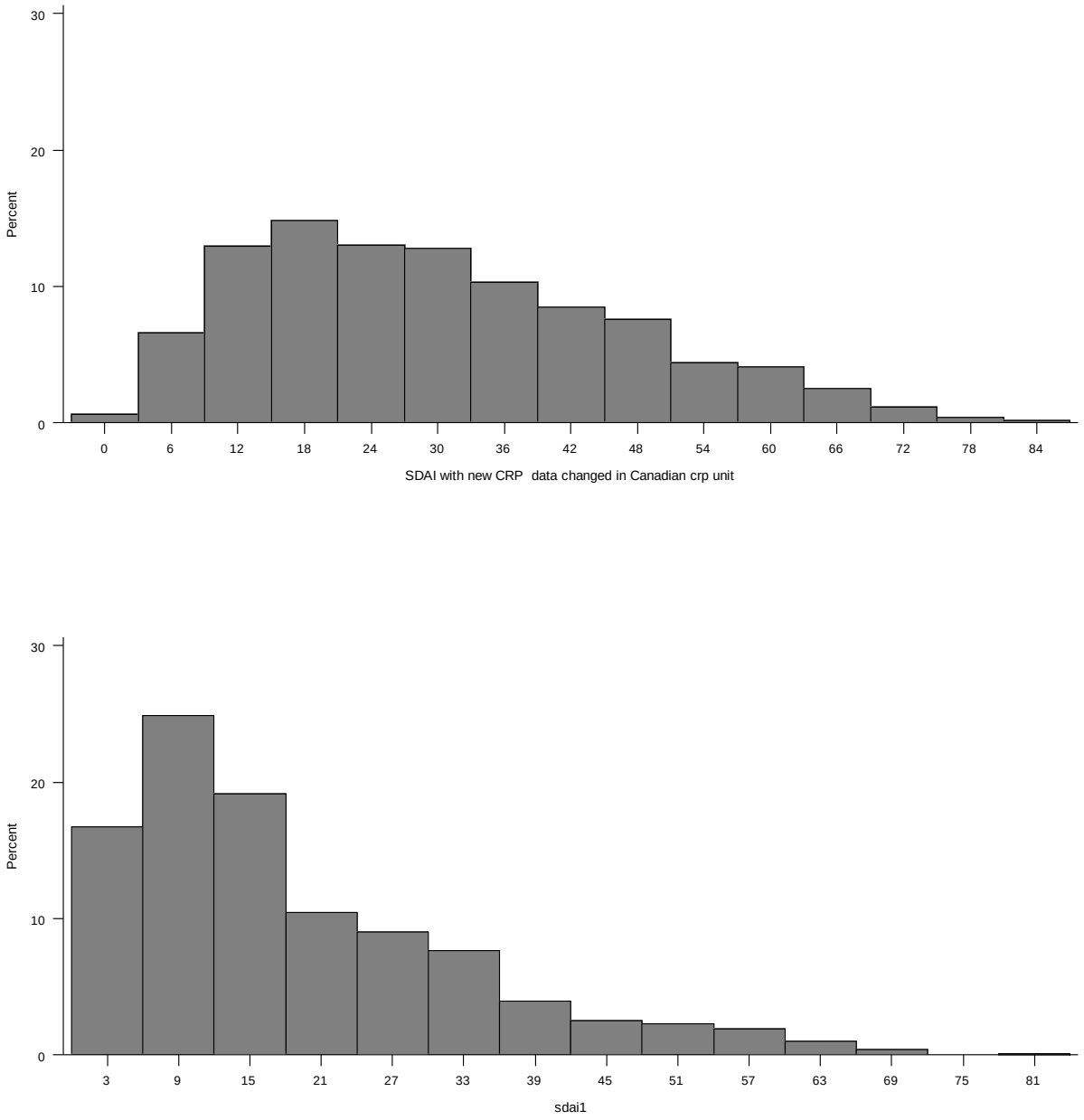

Figure 2- Disease activity (SDAI) distribution at baseline (top) and year 1 (bottom)

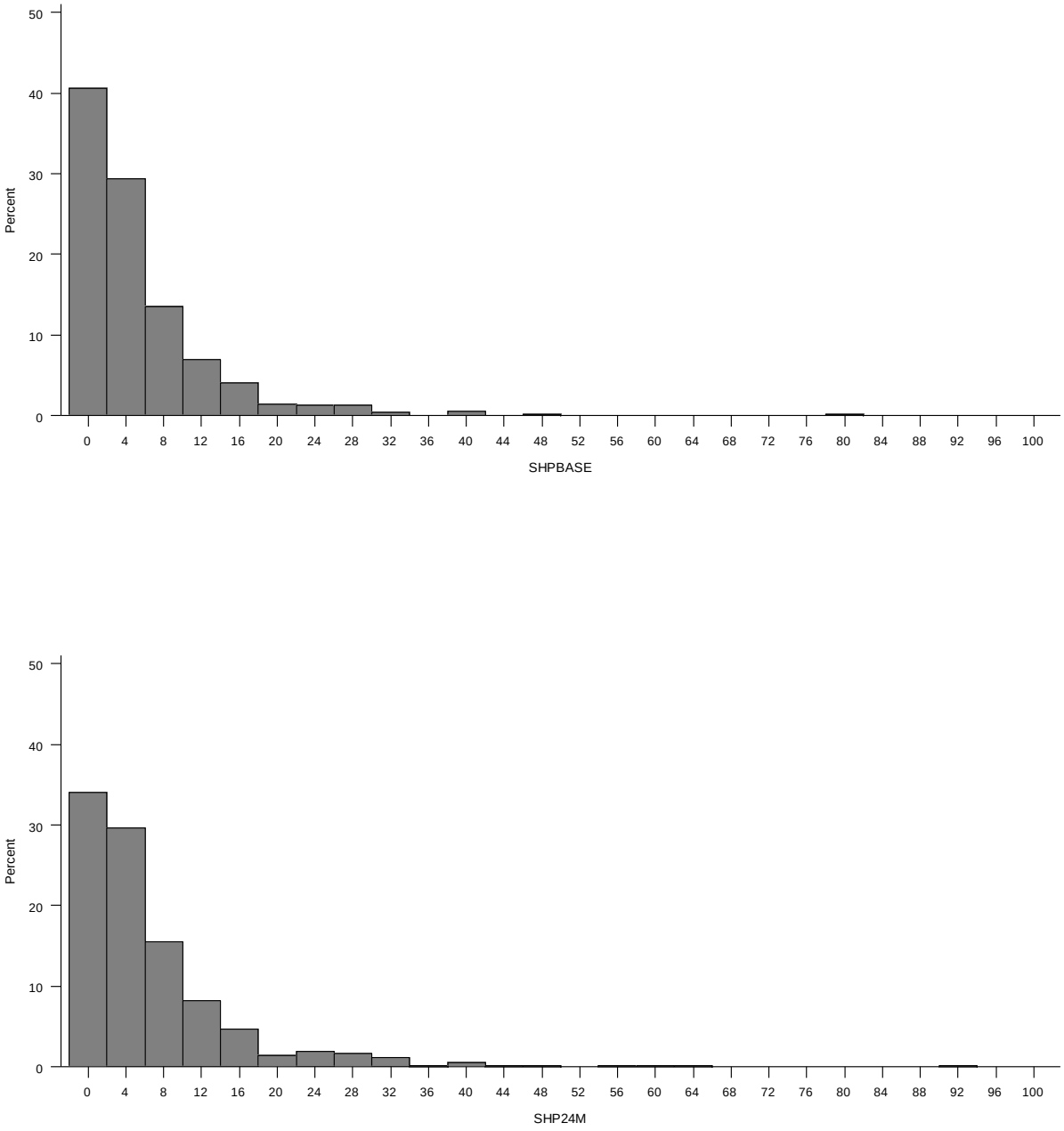

Figure 3- Radiographic damage (Sharp score) distribution at baseline (top) and after 2 years (bottom)

Appendix 2-D

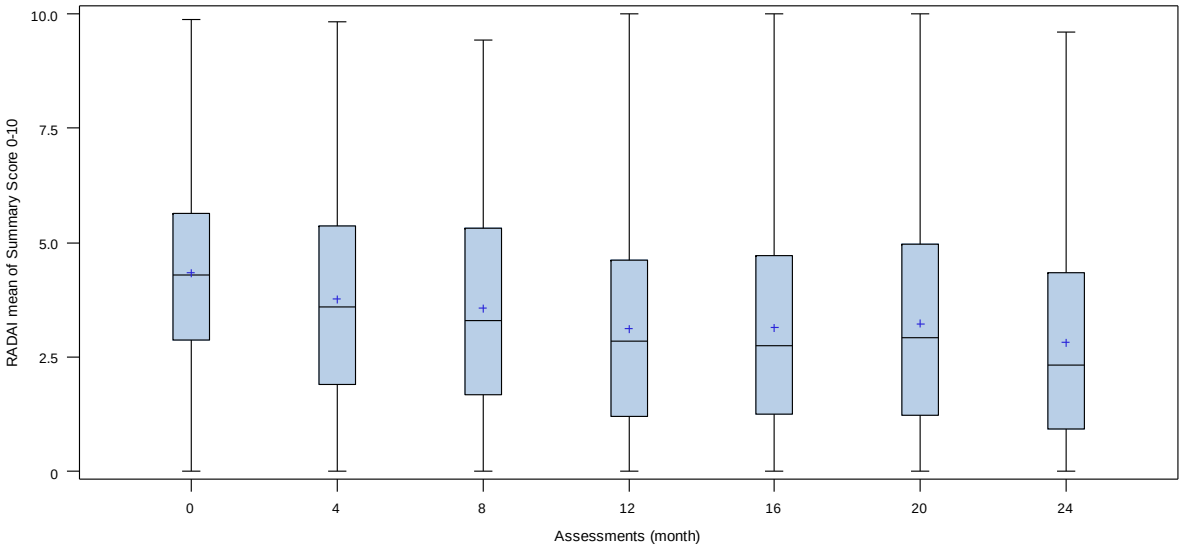

Figure- RADAI improvement pattern over 2 years

Appendix 2-E

Test for multicollinearity . Primary outcome: X-ray progression over 2 years

	VIF	Tolerance
LDAS yr 1	1.15413	0.86645
LDAS baseline	1.10976	0.90109
HAQ baseline	1.18468	0.84411
Female	1.09391	0.91416
Age	1.32266	0.75605
RF	1.90735	0.52429
Anti-CCP	1.93025	0.51807
MTX	1.05289	0.94977
CRP baseline	1.06823	0.93612
SHS baseline	1.34747	0.74213
SHS baseline >0	1.24629	0.80238

Appendix 2-F

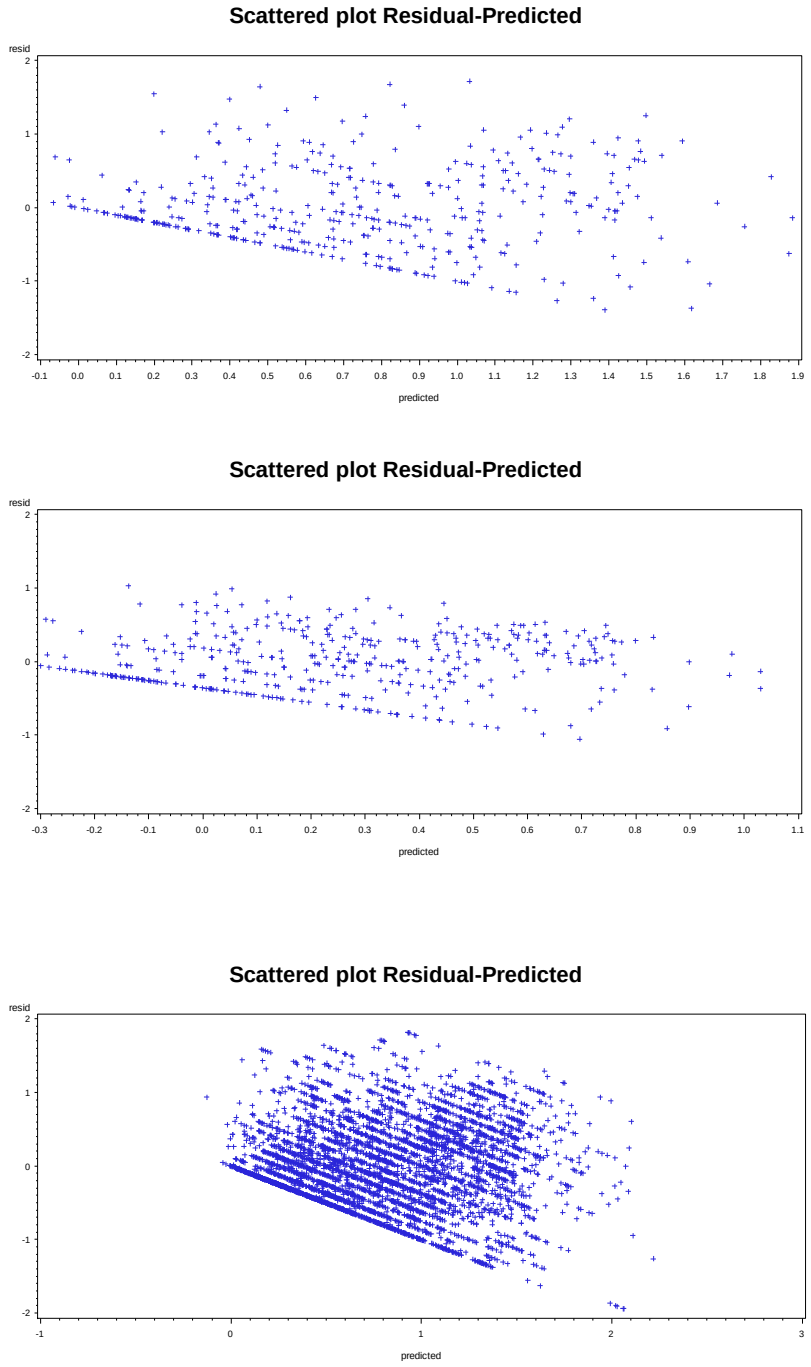

Figure - Scattered plot of Residual-Predicted for outcome HAQ at year 3 (linear regression) before (top) and after (middle) log transformation and after multiple imputation (bottom) based on a similar model (same covariates).

Appendix 2-G

Table 1- Linear regression analysis of complete cases after log transformation with different damage scores in each model. Outcome: HAQ at year 3.

Model including anti-CCP				Model including RF		
R^2= 0.361524				R^2= 0.359743		
Predictor	**β**	**95% CI**	**p-value**	**β**	**95% CI**	**p-value**
LDA yr 1	-0.1298	-0.2093, -0.0503	0.0014	-0.1296	-0.2094, -0.0499	0.0015
LDA BSL	0.0260	-0.0899, 0.1419	0.65	0.0288	-0.0875, 0.1449	0.63
Age	0.0048	0.0019, 0.0077	0.001	0.0047	0.0018, 0.0075	0.0014
Female	0.0945	0.0062, 0.1828	0.036	0.0891	0.0012, 0.1771	0.047
HAQ BSL	0.3373	0.2801, 0.3944	<.0001	0.3390	0.2818, 0.3962	<.0001
TJSN BSL	0.0102	-0.0015, 0.0220	0.09	0.0101	-0.0017, 0.0218	0.09
Anti-CCP	0.0431	-0.0313, 0.1176	0.26			
RF				0.0123	-0.0627, 0.0874	0.75
MTX	0.004	-0.0702, 0.0785	0.91	0.0082	-0.0659, 0.0823	0.83
R^2= 0. 357192				R^2= 0. 355578		
Predictors	**β**	**95% CI**	**p-value**	**β**	**95% CI**	**p-value**
LDA yr 1	-0.1356	-0.2151, -0.0561	0.0009	-0.1354	-0.2151, -0.0557	0.0009
LDA BSL	0.0272	-0.0891, 0.1435	0.65	0.0300	-0.0865, 0.1466	0.61
Age	0.0055	0.0026, 0.0084	0.0002	0.0053	0.0025, 0.0082	0.0003
Female	0.0990	0.0098, 0.1882	0.03	0.0942	0.0053, 0.1831	0.038
HAQ BSL	0.3380	0.2807, 0.3953	<.0001	0.3396	0.2823, 0.3970	<.0001
TERO BSL	0.0005	-0.0082, 0.0092	0.91	0.0008	-0.0078, 0.0095	0.85
Anti-CCP	0.0405	-0.0344, 0.1154	0.29			
RF				0.0092	-0.6610, 0.0845	0.81
MTX	0.0070	-0.0675, 0.0816	0.85	0.0109	-0.0634, 0.0852	0.77
R^2= 0. 358642				R^2= 0. 357118		
Predictors	**β**	**95%CI**	**p-value**	**β**	**95% CI**	**p-value**
LDA yr 1	-0.1332	-0.2127, -0.0537	0.0011	-0.1330	-0.2127, -0.0532	0.0011
LDA BSL	0.0272	-0.0889, 0.1434	0.65	0.0299	-0.0865, 0.1464	0.51
Age	0.0050	0.0021, 0.0080	0.0009	0.0048	0.0019,0.0077	0.0013
Female	0.1024	0.0126, 0.1899	0.025	0.0962	0.0080,0.1845	0.03
HAQ BSL	0.3374	0.2801, 0.3947	<.0001	0.3390	0.0281, 0.3963	<.0001
SHS BSL	0.0032	-0.0031, 0.0094	0.32	0.0033	-0.0030, 0.0096	0.30
Anti-CCP	0.0393	-0.0353, 0.1140	0.30			
RF				0.0092	-0.0659, 0.0844	0.81
MTX	0.0061	-0.0684, 0.0806	0.87	0.0098	-0.0644, 0.0841	0.79

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde; BSL=baseline; TJSN= Total joint space narrowing; TERO= total erosion;

Table 2- Linear regression analysis of imputed data with different damage scores in each model. Outcome: HAQ at year 3.

	Model including anti-CCP			Model including RF		
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA yr 1	-0.2137	-0.0307,-0.1207	<.0001	-0.2120	-0.3048, -0.1192	<.0001
LDA BSL	0.0333	-0.1049, 0.1718	0.63	0.0327	-0.1057, 0.1711	0.64
HAQ BSL	0.4719	0.4180, 0.5259	<.0001	0.4716	0.4177, 0.5256	<0.0001
TJSN BSL	0.0164	0.0021, 0.0308	0.0272	0.0165	0.0022, 0.0308	0.026
Female	0.0930	0.0115, 0.1746	0.0254	0.0932	0.0114, 0.1751	0.0256
Age	0.0053	0.0019, 0.0088	0.0038	0.0054	0.0020, 0.0088	0.0035
MTX	-0.0237	-0.0979, 0.0505	0.5301	-0.0259	-0.0989, 0.0471	0.49
Anti-CCP	0.0045	-0.0747, 0.0838	0.9094			
RF				0.0325	-0.0386, 0.1037	0.37
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA yr 1	-0.2188	-0.3090, -0.1285	<.0001	-0.2173	-0.3076, -0.1270	<.0001
LDA BSL	0.0361	-0.1006, 0.1728	0.60	0.0354	-0.1015, 0.1723	0.61
HAQ BSL	0.4745	0.4200, 0.5291	<.0001	0.4742	0.4196, 0.5289	<.0001
TERO BSL	0.0061	-0.0060, 0.0183	0.29	0.0061	-0.0062, 0.0183	0.30
Female	0.0971	0.0147, 0.1796	0.0211	0.0978	0.0150, 0.1806	0.021
Age	0.0060	0.0033, 0.0087	<.0001	0.0061	0.0034, 0.0088	<.0001
MTX	-0.0143	-0.0895, 0.0609	0.71	-0.0175	-0.0916, 0.0565	0.64
Anti-CCP	0.0046	-0.0861, 0.0770	0.91			
RF				0.0293	-0.0418, 0.1005	0.42
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA yr 1	-0.2142	-0.3022, -0.1262	<.0001	-0.2127	-0.3010, -0.1245	<.0001
LDA BSL	0.0358	-0.1015, 0.2731	0.60	0.0352	-0.1023, 0.1727	0.61
HAQ BSL	0.4706	0.4162, 0.5250	<.0001	0.4703	0.4159, 0.5247	<.0001
SHS BSL	0.0068	0.0005, 0.0132	0.04	0.0068	0.0004, 0.0132	0.04
Female	0.0969	0.0135, 0.1801	0.023	0.0976	0.0139, 0.1812	0.02
Age	0.0051	0.0025, 0.0078	0.0002	0.0052	0.0140, 0.0079	0.0001
MTX	-0.0130	-0.0877, 0.0616	0.73	-0.0162	-0.0897, 0.0574	0.67
Anti-CCP	-0.0032	-0.0867, 0.0801	0.94			
RF				0.0308	-0.0408, 0.1024	0.40

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde score; BSL=baseline; TJSN= Total Joint space narrowing; TERO= total erosion;

Appendix 2-H

Table 1- Logistic regression analysis of complete cases. Outcome: Radiographic damage over 2 years

	Model 1 with anti-CCP			Model 2 with RF		
Predictor	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA yr 1	0.72	0.40-1.27	0.25	0.77	0.44-1.35	0.36
LDA BSL	0.47	0.17-1.28	0.14	0.49	0.18-1.33	0.16
Age	1.01	0.99-1.03	0.31	1.01	0.99-1.03	0.44
Female	0.62	0.36-1.07	0.09	0.56	0.33-0.96	0.04
CRP BSL	1.13	0.90-1.40	0.29	1.12	0.91-1.39	0.29
SHS BSL>0	2.74	1.27-5.91	0.01	2.71	1.26-5.81	0.01
HAQ BSL	0.92	0.63-1.34	0.66	0.95	0.65-1.38	0.78
Anti-CCP	3.49	1.98-6.14	<.0001	-	-	-
RF	-	-	-	2.32	1.32-4.08	0.003
c-statistics	0.735			0.709		
AIC	407.31			419.01*		
Predictors	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA yr 1	0.74	0.42-1.30	0.29	0.78	0.45-1.36	0.39
LDA BSL	0.48	0.18-1.29	0.15	0.50	0.19-1.34	0.17
Age	1.01	0.99-1.03	0.31	1.01	0.99-1.03	0.44
Female	0.60	0.35-1.03	0.06	0.55	0.33-0.94	0.03
CRP BSL	1.13	0.91-1.41	0.27	1.12	0.91-1.39	0.29
SHS BSL>0	2.73	1.27-5.88	0.01	2.69	1.26-5.77	0.01
MTX	0.84	0.50-1.41	0.50	0.94	0.56-1.56	0.80
Anti-CCP	3.54	2.01-6.25	<.0001	-	-	-
RF	-	-	-	2.32	1.32-4.08	0.003
c-statistics	0.736			0.710		
AIC	406.06			419.02		

*Hosmer and Lemeshow Goodness-of-Fit Test p=0.03; MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde score; BSL=baseline;

Table 2: Logistic regression analysis of imputed data. Outcome: Radiographic damage over 2 years

	Model 1 with anti-CCP			Model 2 with RF		
Predictor	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA yr 1	0.61	0.38-0.98	0.04	0.64	0.39-1.39	0.07
LDA BSL	1.02	0.41-2.52	0.96	1.04	0.42-2.57	0.92
Age	1.02	1.01-1.04	0.008	1.02	1.01-1.04	0.004
Female	0.86	0.54-1.36	0.51	0.82	0.52-1.29	0.38
CRP BSL	1.15	1.00-1.32	0.05	1.17	1.02-1.34	0.03
SHS BSL>0	3.38	1.85-6.18	<.0001	3.32	1.82-6.07	0.0001
HAQ BSL	1.04	0.81-1.35	0.74	1.05	0.81-1.36	0.72
Anti-CCP	2.18	1.54-3.07	<.0001	-	-	-
RF	-	-	-	1.90	1.19-3.01	0.009
c-statistics range	0.709- 0.725			0.705- 0.720		
AIC range	982.07- 1035.33			997.56-1050.44		
Predictors	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA yr1	0.60	0.38-0.97	0.04	0.63	0.39-1.02	0.06
LDA BSL	0.99	0.40-2.40	0.97	1.01	0.42-2.45	0.98
Age	1.02	1.01-1.04	0.01	1.02	1.01-1.04	0.00
Female	0.87	0.55-1.37	0.54	0.83	0.52-1.30	0.41
CRP BSL	1.16	1.02-1.32	0.03	1.18	1.03-1.34	0.02
SHS BSL>0	3.37	1.85-6.17	<.0001	3.33	1.82-6.08	0.0001
MTX	0.82	0.52-1.31	0.39	0.87	0.54-1.38	0.52
Anti-CCP	2.24	1.60-3.12	<.0001	-	-	-
RF	-	-	-	1.92	1.19-3.09	0.009
c-statistics range	0.711-0.717			0.704-0.721		
AIC range	982.20-1035.32			997.63-1050.76		

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde score; BSL=baseline;

Appendix 2-I

Table 1: Logistic regression analysis of complete cases. Outcome: Radiographic damage over 2 years. Main predictor LDA (RADAI < 2.2) at 4mo (n=479)

	Model 1 with anti-CCP			Model 2 with RF		
Predictor	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA 4 mo	0.49	0.25-0.97	0.04	0.54	0.28-1.04	0.07
LDA BSL	1.58	0.72-3.46	0.25	1.59	0.74-3.46	0.24
Age	1.01	0.99-1.03	0.19	1.01	0.99-1.03	0.31
Female	0.54	0.31-0.93	0.03	0.49	0.28-0.84	0.01
SHS BSL>0	2.16	1.03-4.52	0.04	2.21	1.06-4.61	0.03
HAQ BSL	1.07	0.71-1.61	0.76	1.11	0.74-1.67	0.61
Anti-CCP	2.89	1.65-5.06	0.0002	-	-	-
RF	-	-	-	1.83	1.05-3.18	0.03
c-statistics	0.713			0.683		
AIC	402.56			412.70		
Predictors	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA 4mo	0.49	0.25-0.95	0.03	0.52	0.27-1.00	0.05
LDA BSL	1.51	0.72-3.13	0.27	1.47	0.72-3.03	0.29
Age	1.01	0.99-1.03	0.15	1.01	0.99-1.03	0.26
Female	0.54	0.31-0.93	0.03	0.50	0.29-0.85	0.01
SHS BSL>0	2.15	1.03-4.51	0.04	2.21	1.06-4.62	0.03
MTX	0.89	0.53-1.49	0.66	0.98	0.59-1.62	0.92
Anti-CCP	2.96	1.68-5.18	0.0002	-	-	-
RF	-	-	-	1.86	1.07-3.22	0.03
c-statistics	0.714			0.684		
AIC	402.45			412.94		

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde score; BSL=baseline; LDA=low disease activity, RADAI= RA disease activity index

Table 2: Logistic regression analysis of complete cases. Outcome: Radiographic damage over 2 years. Main predictor LDA (RADAI < 2.2) at 12mo

	Model 1 with anti-CCP			Model 2 with RF		
Predictor	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA yr 1	0.65	0.36-1.14	0.13	0.69	0.39-1.21	0.19
LDA BSL	1.29	0.61-2.74	0.51	1.33	0.63-2.80	0.46
Age	1.01	0.99-1.03	0.23	1.01	0.99-1.03	0.33
Female	0.64	0.38-1.10	0.11	0.60	0.35-1.02	0.06
SHS BSL>0	2.56	1.24-5.29	0.01	2.52	1.22-5.20	0.01
HAQ BSL	1.10	0.74-1.64	0.63	1.15	0.77-1.70	0.50
Anti-CCP	2.77	1.63-4.73	0.0002	-	-	-
RF	-	-	-	1.93	1.13-3.31	0.02
c-statistics	0.699			0.675		
AIC	433.93			443.03		
Predictors	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA yr 1	0.63	0.36-1.09	0.10	0.66	0.38-1.14	0.13
LDA BSL	1.19	0.59-2.42	0.62	1.20	0.60-2.42	0.60
Age	1.01	0.99-1.03	0.17	1.01	0.99-1.03	0.26
Female	0.65	0.38-1.11	0.11	0.61	0.36-1.04	0.07
SHS BSL>0	2.57	1.25-5.32	0.01	2.55	1.24-5.26	0.01
MTX	0.86	0.52-1.42	0.56	0.95	0.58-1.55	0.84
Anti-CCP	2.86	1.67-4.90	0.0001	-	-	-
RF	-	-	-	1.96	1.15-3.36	0.01
c-statistics	0.700			0.684		
AIC	433.81			443.45		

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ ver der Heijde score; BSL=baseline; LDA=low disease activity, RADAI= RA disease activity index

Appendix 2-J

Table 1: Logistic regression analysis of imputed data. Outcome: Radiographic damage over 2 years. Main predictor LDA (RADAI < 2.2) at 4 mo (n=984)

	Model 1 with anti-CCP			Model 2 with RF		
Predictor	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA 4mo	0.65	0.33-1.28	0.19	0.66	0.34-1.28	0.20
LDA BSL	1.32	0.68-2.60	0.40	1.38	0.70-2.74	0.33
Age	1.02	1.00-1.04	0.01	1.02	1.01-1.04	0.01
Female	0.85	0.58-1.24	0.39	0.81	0.55-1.20	0.29
CRP BSL	1.10	0.98-1.23	0.10	1.10	0.98-1.24	0.10
SHS BSL>0	3.60	1.94-6.61	<.0001	3.53	1.92-6.51	<.0001
HAQ BSL	1.17	0.82-1.66	0.37	1.19	0.85-1.66	0.30
Anti-CCP	1.66	1.10-2.50	0.02	-	-	-
RF	-	-	-	1.72	1.04-2.82	0.04
Predictors	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA 4mo	0.61	0.33-1.50	0.12	0.63	0.34-1.15	0.12
LDA BSL	1.19	0.66-2.14	0.55	1.23	0.68-2.24	0.48
Age	1.02	1.01-1.03	0.00	1.02	1.01-1.04	0.00
Female	0.88	0.60-1.29	0.50	0.84	0.57-1.25	0.39
CRP BSL	1.11	1.00-1.25	0.08	1.11	1.00-1.25	0.07
SHS BSL>0	3.63	1.96-6.73	<.0001	3.58	1.93-6.63	<.0001
MTX	0.91	0.60-1.36	0.63	0.91	0.62-1.34	0.64
Anti-CCP	1.69	1.11-2.57	0.01	-	-	-
RF	-	-	-	1.73	1.05-2.84	0.03

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde score; BSL=baseline; LDA=low disease activity, RADAI= RA disease activity index

Table 2: Logistic regression analysis of imputed data. Outcome: Radiographic damage over 2 years. Main predictor LDA (RADAI < 2.2) at 12mo

	Model 1 with anti-CCP			Model 2 with RF		
Predictor	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA yr 1	0.74	0.50-1.10	0.13	0.75	0.52-1.10	0.14
LDA BSL	1.27	0.67-2.42	0.45	1.34	0.70-2.57	0.37
Age	1.02	1.01-1.03	0.01	1.02	1.01-1.03	0.01
Female	0.85	0.58-1.24	0.39	0.81	0.55-1.20	0.29
CRP BSL	1.11	0.99-1.25	0.06	1.11	0.99-1.24	0.06
SHS BSL>0	3.64	1.98-6.70	<.0001	3.58	1.95-6.57	<.0001
HAQ BSL	1.18	0.84-1.65	0.33	1.19	0.86-1.65	0.27
Anti-CCP	1.65	1.10-2.48	0.02	-	-	-
RF	-	-	-	1.72	1.05-2.81	0.03
Predictors	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
LDA yr 1	0.70	0.49-1.01	0.05	0.71	0.50-1.01	0.06
LDA BSL	1.14	0.65-1.97	0.63	1.19	0.67-2.08	0.54
Age	1.02	1.01-1.03	0.002	1.02	1.009-1.03	0.002
Female	0.88	0.60-1.29	0.49	0.84	0.57-1.25	0.39
CRP BSL	1.12	1.00-1.26	0.04	1.12	1.01-1.26	0.04
SHS BSL>0	3.70	2.00-6.85	<.0001	3.64	1.97-6.72	<.0001
MTX	0.91	0.60-1.37	0.64	0.91	0.62-1.35	0.64
Anti-CCP	1.69	1.11-2.56	0.01	-	-	-
RF	-	-	-	1.74	1.07-2.82	0.03

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde score; BSL=baseline

Appendix 2-K

Table 1. Linear regression analysis of complete case. Outcome: HAQ at year 3; Main predictor: LDA at 4 month (RADAI <2.2) (n=449)

Predictor	β	95%CI	p-value	β	95%CI	p-value
LDA 4mo	-0.2194	-0.3466,-0.0923	0.0008	-0.2148	-0.3421, -0.0875	0.001
LDA0	0.0269	-0.1404, 0.1941	0.75	0.0299	-0.1380, 0.1979	0.73
HAQ BSL	0.4962	0.4109, 0.5815	<.0001	0.4995	0.4141, 0.5849	<0.0001
TJSN BSL	0.0135	-0.0033, 0.0304	0.12	0.0132	-0.0037, 0.0301	0.13
Female	0.1327	-0.0076, 0.2454	0.06	0.1095	0.0165, 0.2355	0.09
Age	0.0082	0.0041, 0.0126	<.0001	0.0081	0.0039, 0.0122	0.0001
MTX	0.0071	-0.0984, 0.1127	0.89	0.0136	-0.0917, 0.1190	0.80
Anti-CCP	0.0836	-0.0223, 0.1896	0.12			
RF				0.0405	-0.0667, 0.1477	0.46
R2	0.3740			0.3714		
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA 4mo	-0.2258	-0.3531, -0.0986	0.0005	-0.2213	-0.3486, -0.0939	0.0007
LDA0	0.0260	-0.1420, 0.1940	0.76	0.0294	-0.1392, 0.1981	0.73
HAQ BSL	0.4974	0.4118, 0.5830	<.0001	0.5006	0.4149, 0.5862	<.0001
TERO BSL	-0.0069	-0.0132, 0.0118	0.91	-0.0004	-0.0128, 0.01219	0.96
Female	0.1224	0.0055, 0.2502	0.06	0.1137	-0.037, 0.2411	0.08
Age	0.0092	0.0052, 0.0134	<.0001	0.0091	0.0050, 0.0132	<.0001
MTX	0.0105	-0.0952, 0.1163	0.84	0.0167	-0.0888, 0.1223	0.75
Anti-CCP	0.0803	-0.0260, 0.1866	0.14			
RF				0.0378	-0.0697, 0.1453	0.49
R2	0.3705			0.3680		
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDAS 4mo	-0.2246	-0.3354, -0.1306	0.0006	-0.2200	-0.3436, -0.0927	0.0007
LDAS0	0.0293	-0.1205, 0.1205	0 73	0.0325	-0.1359, 0.2009	0.70
HAQ BSL	0.4964	0.4097, 0.5358	<.0001	0.4996	0.4140, 0.5852	<.0001
Sharp BSL	0.0036	-0.0056, 0.0127	0.44	0.0037	-0.0055, 0.0128	0.43
Female	0.1270	-0.0000, 0.2539	0.05	0.1180	-0.0085, 0.2325	0.07
Age	0.0086	0.0044, 0.0128	<.0001	0.0084	0.0042, 0.0126	0.0001
MTX	0.0099	-0.0958, 0.1156	0.85	0.0160	-0.0895, 0.1215	0.77
Anti-CCP	0.0792	-0.0269, 0.1853	0.14			
RF				0.0375	-0.0699, 0.1448	0.49
R2	0.3713			0.3689		

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde score; BSL=baseline; TJSN= Total Joint space narrowing; TERO= total erosion

Table 2. Linear regression analysis of complete case. Outcome: HAQ at year 3; Main predictor: LDA at 12 month (RADAI <2.2) (n=449)

Predictor	β	95%CI	p-value	β	95%CI	p-value
LDA yr 1	-0.3248	-0.4384, -0.2112	<.0001	-0.3245	-0.4383, -0.2107	<.0001
LDA BSL	0.0549	-0.1078, 0.2175	0.51	0.0584	-0.1046, 0.2216	0.48
HAQ BSL	0.4690	0.3849, 0.5530	<.0001	0.0471	0.3871, 0.5554	<0.0001
TJSN BSL	0.0134	-0.0031, 0.0299	0.11	0.0131	-0.0034, 0.0296	0.12
Female	0.1162	-0.0075, 0.2400	0.06	0.1079	0.0154, 0.2311	0.09
Age	0.0085	0.0045, 0.0125	<.0001	0.0084	0.0044, 0.0124	<.0.0001
MTX	0.0145	-0.0888, 0.1178	0.78	0.0202	-0.0828, 0.1234	0.70
Anti-CCP	0.0756	-0.0279, 0.1792	0.15			
RF				0.0392	-0.0656, 0.1441	0.46
R2	0.4007			0.3986		
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA yr 1	-0.3288	-0.4427, -0.2149	<.0001	-0.3285	-0.4426, -0.2144	<.0001
LDA BSL	0.0539	-0.1094, 0.2173	0.51	0.0578	-0.1061, 0.2216	0.50
HAQ BSL	0.4702	0.3859, 0.5546	<.0001	0.4724	0.3880, 0.5568	<.0001
TERO BSL	0.0002	-0.0120, 0.0125	0.97	0.0006	-0.0116, 0.0128	0.92
Female	0.1210	-0.0041, 0.2461	0.06	0.1133	-0.0113, 0.2380	0.07
Age	0.0094	0.0055, 0.0135	<.0001	0.0093	0.0052, 0.0133	<.0001
MTX	0.0180	-0.0855, 0.1215	0.73	0.0235	-0.07981, 0.1268	0.65
Anti-CCP	0.0716	-0.0323, 0.1756	0.18			
RF				0.0363	-0.0689, 0.1414	0.50
R2	0.3972			0.3953		
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA yr 1	-0.3288	-0.4426, -0.2151	<.0001	-0.3285	-0.4424, -0.2145	<.0001
LDA BSL	0.0574	-0.1058, 0.2205	0.49	0.0610	-0.1026, 0.2246	0.46
HAQ BSL	0.4691	0.3849, 0.5533	<.0001	0.4712	0.3869, 0.5556	<.0001
Sharp BSL	0.0040	-0.0049, 0.0130	0.37	0.0041	-0.0048, 0.0131	0.36
Female	0.1247	0.0005, 0.2490	0.05	0.1168	-0.0069, 0.2405	0.06
Age	0.0088	0.0047, 0.0130	<.0001	0.0087	0.0045, 0.0128	<.0001
MTX	0.0172	-0.0862, 0.1207	0.74	0.0226	-0.0806, 0.1258	0.67
Anti-CCP	0.0710	-0.0327, 0.1747	0.18			
RF				0.0534	-0.0688, 0.1412	050
R2	0.3983			0.3964		

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde; BSL=baseline; TJSN= Total joint space narrowing; TERO= total erosion; RADAI=RA disease activity index

Appendix 2-L

Table 1. Linear regression analysis of imputed data. Outcome: HAQ at year 3; Main predictor: LDA at 4 month (RADAI <2.2)

Predictor	β	95%CI	p-value	β	95%CI	p-value
LDA 4mo	-0.2252	-0.3329,-0.11075	<.0001	-0.2205	-0.3279, -0.1131	<.0001
LDA BSL	0.0221	-0.1166, 0.1609	0.75	0.0252	-0.1136, 0.1640	0.72
HAQ BSL	0.4828	0.4124, 0.5533	<.0001	0.4848	0.4146, 0.5550	<0.0001
TJSN BSL	0.0172	0.0038, 0.0307	0.01	0.0168	0.0033, 0.0302	0.014
Female	0.1327	0.0302, 0.2352	0.01	0.1287	0.0263, 0.2311	0.014
Age	0.0081	0.0047, 0.0115	<.0001	0.0082	0.0047, 0.0115	<.0001
MTX	-0.0382	-0.1254, 0.0490	0.39	-0.0353	-0.1221, 0.0515	0.492
Anti-CCP	0.0754	-0.0250, 0.1758	0.14			
RF				0.0533	-0.0345, 0.1412	0.23
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA 4mo	-0.2284	-0.3390, -0.1179	<.0001	-0.2223	-0.3324, -0.1123	<.0001
LDA BSL	0.0527	-0.0845, 0.1899	0.45	0.0567	-0.0804, 0.1939	0.42
HAQ BSL	0.4815	0.4089, 0.5541	<.0001	0.4843	0.4120, 0.5566	<.0001
TERO BSL	0.0063	-0.0032, 0.0159	0.19	0.0065	-0.0030, 0.0161	0.18
Female	0.1417	0.0380, 0.2455	0.007	0.1379	0.0343, 0.2416	0.009
Age	0.0087	0.0053, 0.0121	<.0001	0.0086	0.0052, 0.0121	<.0001
MTX	-0.0465	-0.1349, 0.0417	0.30	-0.0440	-0.1320, 0.0439	0.32
Anti-CCP	0.0775	-0.0245, 0.1796	0.13			
RF				0.0620	-0.0271, 0.1512	0.17
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA 4mo	-0.2330	-0.3354, -0.1306	<.0001	-0.2304	-0.3327, -0.1281	<.0001
LDA BSL	-0.0050	-0.1205, 0.1205	0.94	0.0032	-0.1288, 0.1224	0.96
HAQ BSL	0.4747	0.4097, 0.5358	<.0001	0.4737	0.4106, 0.5368	<.0001
SHS BSL	0.0026	-0.0031, 0.0083	0.35	0.0026	-0.0032, 0.0083	0.36
Female	0.1435	0.0523, 0.2343	0.002	0.1420	0.0514, 0.2325	0.002
Age	0.0083	0.0052, 0.0114	<.0001	0.0084	0.0053, 0.0114	<.0001
MTX	-0.0260	-0.1045, 0.0526	0.52	-0.0264	-0.1048, 0.0520	0.51
Anti-CCP	0.0283	-0.0629, 0.1196	0.54			
RF				0.0404	-0.0395, 0.1203	0.32

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde; BSL=baseline; TJSN= Total joint space narrowing; TERO= total erosion; RADAI=RA disease activity index

Table 2. Linear regression analysis of imputed data. Outcome: HAQ at year 3; Main predictor: LDA at 12 month (RADAI <2.2)

	Model with anti-CCP			Model with RF		
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA yr 1	-0.3640	-0.4564,-0.2717	<.0001	-0.3630	-0.4553, -0.2707	<.0001
LDA BSL	0.0595	-0.0741, 0.1931	0.38	0.0631	-0.0703, 0.1965	0.35
HAQ BSL	0.4469	0.3782, 0.5156	<.0001	0.4482	0.3796, 0.5168	<0.0001
TJSN BSL	0.0168	0.0038, 0.0299	0.01	0.0164	0.0033, 0.0294	0.014
Female	0.1316	0.0038, 0.2310	0.009	0.1278	0.0285, 0.2270	0.012
Age	0.0088	0.0055, 0.0121	<.0001	0.0088	0.0055, 0.0121	<.0001
MTX	-0.0314	-0.1157, 0.0528	0.46	-0.0292	-0.1133, 0.0549	0.49
Anti-CCP	0.0736	-0.0227, 0.1698	0.13			
RF				0.0589	-0.0262, 0.1441	0.17
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA yr1	-0.3540	-0.4472, -0.2609	<.0001	-0.3523	-0.4455, -0.2593	<.0001
LDA BSL	0.0815	-0.0505, 0.2136	0.23	0.0862	-0.0455, 0.2179	0.20
HAQ BSL	0.4529	0.3825, 0.5232	<.0001	0.4547	0.3845, 0.5248	<.0001
TERO BSL	0.0052	-0.0040, 0.0145	0.27	0.0055	-0.0038, 0.0148	0.25
Female	0.1329	0.0321, 0.2338	0.01	0.1296	0.0289, 0.2303	0.01
Age	0.0093	0.0060, 0.0126	<.0001	0.0093	0.0059, 0.0126	<.0001
MTX	-0.0143	-0.0251, 0.1691	0.38	-0.0363	-0.1217, 0.0491	0.40
Anti-CCP	0.0046	-0.1236, 0.0475	0.14			
RF				0.0627	-0.0238, 0.1491	0.15
Predictor	**β**	**95%CI**	**p-value**	**β**	**95%CI**	**p-value**
LDA yr1	-0.3981	-0.4807, -0.3154	<.0001	-0.3975	-0.4801, -0.3149	<.0001
LDA BSL	0.0393	-0.0799, -0.1586	0.52	0.0420	-0.0770, 0.1611	0.49
HAQ BSL	0.4393	0.3789, 0. 4997	<.0001	0.4401	0.3797, 0.5005	<.0001
SHS BSL	0.0021	-0.0035, 0.0078	0.43	0.0021	-0.0035, 0.0077	0.44
Female	0.1420	0.0547, 0.2293	0.001	0.1404	0.0533, 0.2275	0.002
Age	0.0092	0.0063, 0.0122	<.0001	0.0093	0.0063, 0.0122	<.0001
MTX	-0.0180	-0.0936, 0.0575	0.64	-0.0188	-0.0942, 0.0566	0.62
Anti-CCP	-0.0332	-0.0505, 0.1169	0.44			
RF				0.0494	-0.0274, 0.1262	0.21

MTX=methotrexate; RF=Rheumatoid factor; SHS=modified sharp/ver der Heijde; BSL=baseline; TJSN= Total joint space narrowing; TERO= total erosion; RADAI=RA disease activity index