Individuals at risk of Rheumatoid Arthritis: Evaluating Phenotypes and Predictors of Disease

Laura Elizabeth Hunt

Submitted in accordance with the requirements for the degree of
Doctor of Medicine (MD)

The University of Leeds

Faculty of Medicine and Health

Leeds Institute of Rheumatic and Musculoskeletal Medicine

November 2017

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The background literature review in Chapter 2 is based on a review article for Nature Reviews Rheumatology, for which the candidate was first author. The candidate performed the literature review, appraised articles and wrote the manuscript. Senior author (PE) reviewed and provided editorial direction.

Chapter 6 is based on work from a peer reviewed publication. The candidate is first author and recruited the patients, collated the clinical data, analytically appraised literature, analysed data, and wrote the manuscript. The co-authors roles were as follows: EMH, provided statistical analysis. JN, designed the clinical study, contributed to patients' data acquisition. ANB, contributed to laboratory data acquisition, review of the manuscript. RP, contributed to laboratory data acquisition. PE designed the clinical study, and reviewed the manuscript. FP, designed the laboratory data study, analysed the data and wrote and reviewed the manuscript.

The publications are as follows:

Hunt L, and Emery P, Defining populations at risk of rheumatoid arthritis: the first steps to prevention. Nature Reviews Rheumatology, 2014: 10(9):521-30

Hunt L, Hensor E M, Nam J, Burska A N, Parmar R, Emery P, Ponchel F. T cell subsets: an immunological biomarker to predict progression to clinical arthritis in ACPA-positive individuals. Annals of the Rheumatic Diseases, 2016: 75(10):1884-9

Acknowledgements

The work conducted in this thesis has been possible thanks to the supportive network of colleagues from the Leeds Institute of Rheumatic and Musculoskeletal Medicine at Chapel Allerton Hospital. My programme of work has been supported by the infrastructure and resources funded by the National Institute for Health Research Biomedical Research Centre.

I would like to thank the support and guidance of my supervisors Professor Paul Emery, Professor Maya Buch and Dr Sinisa Savic. Particular thanks are given to Paul and Maya who provided both clinical and educational support, assisting my development as a researcher and a clinician.

This thesis focusses on individuals recruited to the CCP study. I was responsible for co-ordinating this research clinic, patient recruitment and data collection. I have worked with several clinical research fellows during this period and wish to acknowledge the input of Drs Kulveer Mankia and Chadi Rakieh. I would like to give particular thanks to the guidance provided by Dr Jaqueline Nam, who set up and coordinated the CCP study through its infancy. Jackie has provided invaluable experience and knowledge of the cohort and performed the ultrasound scans. The administrative, laboratory, nursing support of the CCP research clinic created an effective research environment. I would like to thank all members of the team for their input, including, the research nurses; David Pickles and Kate Russell, the lab team; Diane Corscadden and Katie Mbara and the clinical trial assistants; Ian Weatherill, Christopher Brookes and Phil Luxford.

This research has involved working with a number of specialist. I would like to acknowledge the work of the radiologists' Drs Andrew Grainger and Emma Rowbotham, who scored the MR images reported in Chapter 4. Chapter 5 involved working closely with a post doctorate colleague, Dr Lylia Ouboussad. I would like to thank her for the time and guidance she provided regarding miRNA profiling. Both Lylia and I contributed to the analytical process and have co-authored the manuscript which is now in press. I was fortunate to work with Dr Frederique Ponchel for the T-cell subset project - forming

the content of chapter 6. Members of her team (Drs Rekha Parmar and Agata Burska) processed the samples for T-cell subset quantification and flow cytometry gating strategy was performed by Frederique Ponchel. I was responsible for identification of samples, data collection and interpretation. For each of the results chapters, statistical analysis and oversight was provided by the departmental biomedical statistician, Dr Elizabeth Hensor. I would like to thank her for the guidance and training she provided in statistical modelling.

Lastly, I would like to thank my family and friends, especially my husband and daughter, for all their encouragement and support throughout this period of research.

List of publications/presentations arising from this thesis

Original Articles

Ouboussad, L*, **Hunt, L***,Hensor, E, Nam, J, Barnes, N, Emery, P, McDermott, M, Buch, M. Profiling microRNAs in individuals at-risk of progression to Rheumatoid Arthritis. *In press* Arthritis Research and Therapy

Nam J, Hensor E M, **Hunt L**, Conaghan P G, Wakefield R J, Emery P. Ultrasound findings predict progression to inflammatory arthritis in anti-CCP antibody-positive patients without clinical synovitis Ann Rheum Dis 2016 Dec;75(12):2060-2067.

Hunt L, Hensor E M, Nam J, Burska A N, Parmar R, Emery P, Ponchel F. T cell subsets: an immunological biomarker to predict progression to clinical arthritis in ACPA-positive individuals Ann Rheum Dis 2016 Oct;75(10):1884-9.

Nam JL, **Hunt L**, Hensor EM, Emery. Enriching case selection for imminent RA: the use of anti-CCP antibodies in individuals with new non-specific musculoskeletal symptoms - a cohort study. Ann Rheum Dis 2016 Aug;75(8):1452-6.

Rakeih C, Nam J, **Hunt L.** Hensor, E. M, Das, S., Bissell, L. A., Villeneuve, E., McGonagle, D., Hodgson, R., Grainger, A., Wakefield, R. J., Conaghan, P. G., Emery, P. Predicting the development of clinical arthritis in anti-CCP positive individuals with non-specific musculoskeletal symptoms: a prospective observational cohort study. Ann Rheum Dis 2015 Sep;74(9):1659-66.

Review articles

Hunt L, Emery P. Defining populations at risk of rheumatoid arthritis - The first steps to prevention. Nat Rev Rheumatol 2014 Sep;10(9):521-30.

Burska A, **Hunt L**, Boissinot M, Strollo R, Ryan B.J, Vital E.M, Nissim A, Winyard P.G, Emery P and Ponchel F. Autoantibodies to post-translational modifications in Rheumatoid Arthritis. Mediators of Inflammation, 2014;2014:492873.

Hunt L, Emery P. Etanercept in the treatment of rheumatoid arthritis. Expert Opin Biol Ther. 2013 Oct;13(10):1441-50.

Hunt L, Buch M. The 'therapeutic window' and treating to target in rheumatoid arthritis. Clin Med. 2013 Aug;13(4):387-90.

Book Chapter

Hunt L, Eugénio G, Grainger AJ. Magnetic resonance imaging in individuals at risk of rheumatoid arthritis *In Press.* Best Practice & Research Clinical Rheumatology

Oral presentations

Hunt L. Individuals 'at-risk' - Leeds Cohort. Presented at Meeting of Study Group for Risk factors for RA at European Workshop for Rheumatology Research, York, UK 2016.

K Mankia, **L Hunt,** E Hensor, J Nam, V Clerehugh, A Speirs, A Tugnait, T Do, J Meade, D Devine, P Emery. Increased Prevalence of Periodontal Disease in anti-CCP Positive Individuals At Risk of Progression to Inflammatory Arthritis: A Target for Prevention? Presented at European League against Rheumatism annual meeting, London UK 2016. Presented by L Hunt on behalf of K Mankia.

L. Ouboussad*, L. Hunt*, E.M. Hensor, C. Wong, P. Emery, MF McDermott, MH Buch. Identification of A Predictive MicroRNA Signature of Progression from Pre-RA Systemic Autoimmunity to Development of RA Using Matched Serum Samples Presented at European League against Rheumatism annual meeting, Rome, Italy 2015. Presented by L Ouboussad.

* denotes joint first authorship

Abstract

Clinical outcomes in patients with rheumatoid arthritis (RA) are substantially improved by early therapeutic intervention; however many individuals still develop co-morbidities with a significant socio-economic costs. In other conditions, there is a focus on disease prevention. Identification of the earliest signs of disease, or even those at risk of disease, remains challenging due to the heterogeneity of presentations and the complexities of RA pathogenesis. In working towards RA prevention an understanding of the pathology prior to clinical disease is required. If risk of subsequent disease could be accurately quantified, the opportunity to intervene with therapy that might delay or even prevent disease becomes feasible.

This thesis outlines a programme of work primarily focusing on individuals with systemic autoimmunity, but no synovitis. By studying an 'at-risk' group, the stages and phenotypes prior to disease can be described. Clinical, imaging, molecular and cellular biomarkers will be considered in an attempt to characterise individuals risk and assist in the prediction of RA development.

Table of Contents

List of Tables

List of Figures

List of Boxes

1. Introduction

1.1 Background

Rheumatoid arthritis (RA) is an autoimmune disease characterised by a chronic inflammatory arthritis (IA). Inadequately treated, it results in joint destruction with subsequent deformity, disability and substantial socio-economic costs. However, RA outcomes are significantly improved if treated at the earliest point or 'window of opportunity' [1-5]. This concept draws parallels with cancer management, where early initiation of therapy equates to less disease burden and leads to maximum effect of therapy [1]. In RA, this translates to suppressing the inflammation before the irreversible joint damage occurs. Advancement in chronic disease management has explored the concept of arresting or reversing disease. Examples from cardiovascular disease (CVD) demonstrate how targeting 'at-risk' individuals aims to prevent future disease [6].

Prevention of RA is a novel approach which is still in its infancy. It is acknowledged, that the complexities of RA pathogenesis need to be better understood for prevention to be achieved. Data from observational cohorts allows an understanding of the factors involved in the different phases of RA development. By establishing the risks and initiators of disease, insights on when and how to intervene can be developed.

This thesis focuses on a cohort of individuals defined as being 'at risk' of RA. This was defined on the basis of the presence of non-specific musculoskeletal symptoms plus systemic autoimmunity, specifically, the presence of anti-citrullinated protein antibodies (ACPAs). Individuals were recruited into a prospective observation cohort from 2007 and subsequently followed-up during which time a proportion developed the signs and symptoms of clinical inflammatory arthritis. Markers associated with disease progression to inflammatory arthritis were identified (Figure 1).

1.2 Structure of Thesis

Chapter Two: *Literature review*

Following a brief review of the pathogenesis, natural history and prognosis of RA, this chapter focuses on the populations 'at risk' of RA. It aims to review what has been established in these cohorts and the concept of risk stratification and early intervention.

Chapter Three: *Methods: Study Design and Population*

The study design adopted to recruit, investigate and observe the study population is described here, including methods for acquiring the clinical and imaging data. Methods relating to specific biomarkers e.g. microarray analysis, flow cytometry, are described in the appropriate chapter (Chapters five and six respectively).

Chapter Four: *Magnetic Resonance Imaging in individuals with systemic autoimmunity and arthralgia: MRI as an imaging biomarker*

In this chapter, the MRI characteristics of the at-risk cohort are reported. The chapter aims to consider whether MR imaging findings can predict future progression to RA.

Chapter Five: *MiRNA profiling of matched samples in individuals with systemic autoimmunity and arthralgia who progress to RA*

In this chapter, microRNA profiling of at-risk individuals is reported. It aims to identify a signature which is associated with progression to RA.

Chapter Six: *T-cell subsets in individuals with systemic autoimmunity and arthralgia: an immunological biomarker*

In this chapter, the immunological phenotype (through T-cell subset quantification) of at-risk individuals is reported. Associations to progression to RA and potential clinical utility of T-cell subsets as a biomarker is discussed.

Chapter Seven: *Concluding remarks and future directions*

The final chapter is a summation of the conclusions drawn from each chapter of this programme of work. Recent developments within the field are explored as well as the potential avenues for future work.

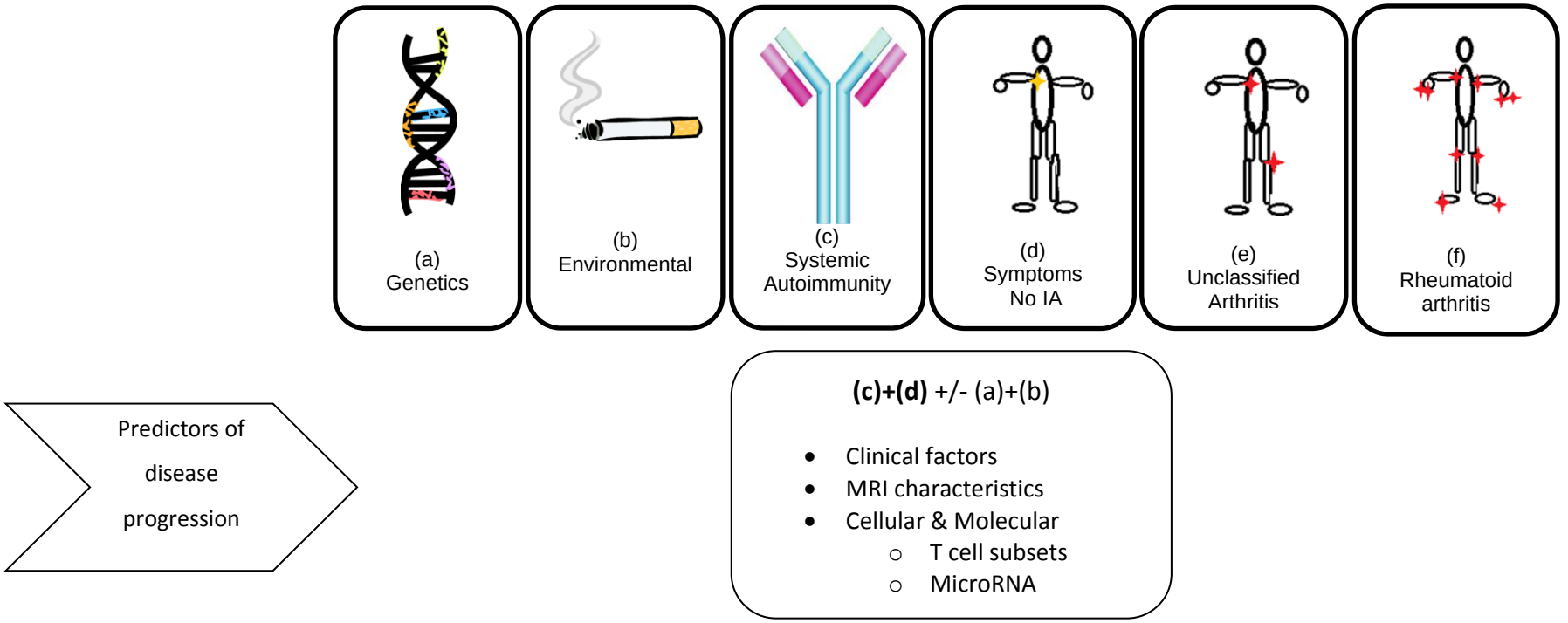


Figure 1 Schematic illustrating inflammatory arthritis continuum with at-risk terminology and highlighting cohort characteristics studied in this thesis.

Presented are the categorises of risk as defined by the EULAR at-risk task force [153]. The central box defines the categorises of individuals that were considered in this programme of work. Individuals with systemic autoimmunity and symptoms were recruited. These individuals may also have genetic factors and environmental exposures.

2. Literature Review

Review Criteria

This review considers the pathogenesis, natural history, diagnosis and prognosis of RA; specific focus is given to the development of RA in at-risk individuals and the strategies to consider for disease prevention as a future approach.

The Cochrane library, Embase and Medline databases were searched using keywords and subject headings including rheumatoid arthritis, secondary prevention, primary prevention, risk factors, arthralgia, preclinical arthritis, biological markers, autoantibodies, develop*/prevent*/predict* arthritis. The search included articles published from 1946 to the present and was restricted to human subjects and the English language. Additional data of interest was sourced from published abstracts and articles. Data generated from this review has been published [7]. Subsequently, articles of interested have been sourced to update the review.

2.1 Rheumatoid Arthritis

RA is a chronic inflammatory autoimmune disease affecting the medium and small joints in a symmetrical pattern. The synovium is the primary site of the characteristic inflammatory pannus formed from the invading immune cells (T cells, macrophages and B cells). This synovial inflammation, or synovitis, results in pain, swelling and stiffness observed in active RA joints. Tissue destruction of cartilage, bone marrow and connective tissues results as the pannus spreads to the articular surface and erodes the bone with the assistance of osteoclasts. The resultant joint damage has been shown to correlate with disability, particularly in late disease. Radiographic damage can be present from as early as 3 months from disease onset [8, 9]. Furthermore, the dysregulated immune system and inflammatory milieu can cause systemic effects on blood vessels and organs, including those in the cardiovascular and respiratory systems [10]. Systemic inflammation contributes to an increased cardiovascular risk, which is at a magnitude comparable to diabetes mellitus [11]. If sub-optimally controlled, RA causes joint damage, deformity, functional impairment and disability [12, 13]. It is

recognised that limiting disease burden and inflammation improves outcome and maintains function. Considerable advances have been achieved in the last 25 years with early identification of disease, optimal treatment strategies and immunotherapies such as the biologic agents [14]. As such, an individual diagnosed today is likely to have a more favourable outcome than in the past and is likely to achieve a state of low disease activity or even remission [15].

2.1.1 Epidemiology

RA affects approximately 0.5-1% of the population with 5-50/100,000 people developing RA each year. There is considerable geographical variation [16]. Northern European countries have higher prevalence and incidence rates compared to southern Europe, whilst north America has the highest overall reported rates [17]. In particular, a high occurrence of RA is reported in the native American-Indian populations; 5.3% in the Pima Indians [18] and 6.8% Chippewa Indians [19]. A UK study reported RA prevalence at 1.16% in females and 0.44% in males, which suggested a decline in incidence in females since the 1950s [20, 21]. This decline has also been noted in the United States [22]. The recently published classification criteria for RA and treat to target approach in disease management may well be responsible for some of these observations.

2.1.2 Aetiology & Pathogenesis

Despite progress in the understanding of specific pathways implicated in established disease, the exact aetiology of RA remains unknown. The current hypothesis describes how individuals with genetic susceptibility interact with environmental determinants resulting in an adverse immune state. A 'triggering event' subsequently leads to T-cell activation, loss of systemic tolerance and the presentation of disease [23]. Identifying which trigger might evoke disease has proved complex due to the multifactorial nature of disease initiation. Recently, the concept of localised mucosal prior to systemic autoimmunity has been introduced [24]. This is primarily attributed to the interaction of environmental factors at various mucosal sites. There are however several caveats to the current hypothesis including whether all individuals require a mucosal trigger prior to systemic disease or whether a break in immune tolerance is possible through

this mechanism. Therefore, much remains to be elucidated to complete our understanding of disease initiation. Specific aetiological factors relating to this thesis will be discussed in this section.

2.1.2.1 Autoantibodies

The presence of autoantibodies in the sera of patients, although not specific for RA, is one of the hall marks of the disease. Autoantibodies targeting the Fc portion of human IgG, termed as 'rheumatoid factor' (RF), were the first to be discovered in the blood of affected individuals [25, 26]. Since its discovery, RF has been consistently reported as a marker of disease progression and radiographic damage [27, 28]. High titres of RF have also been associated with reduced function and disability [29].

IgM RF (but also IgA and IgG RF) in the joint form immune complexes which are thought to initiate a complement cascade. The resultant increased vascular permeability and attraction of chemotactic factors facilitates the influx of immune cells that are integral to the adaptive and innate immune pathways. Their presence alone is not sufficient to initiate disease as evidenced by their detection in healthy individuals. However, at high titre they appear to be more specific for disease [29, 30]. Nell *et al* report high sensitivity and specificity for RA diagnosis using the optimum RF cut of ≥50 U/ml [30].

Anti-citrullinated peptide antibodies (ACPAs) refer to antibodies directed against multiple citrullinated proteins (e.g. fibrinogen, vimentin, and α-enolase) and potentially have a more pathogenic role. These were first discovered in 1960s under the guise of antiperinuclear and antikeratin factor, later shown to be antibodies against a similar antigen, citrulline [31-33]. Citrullination is a post-translational modification of arginine to citrullline which occurs in the presence of high calcium by the enzyme peptidylarginine deiminase. It is a physiological process and occurs in health as well as disease states. The second generation cyclic citrullinated peptide (anti-CCP 2) test is the most frequently used assay to detect ACPA in clinical practice.

ACPAs can be detected in the sera of individuals destined to develop RA years before there is any evidence of inflammation and immunity in the joints [34]. Before the onset of disease, the epitope recognition profile and isotype usage of the ACPA increases implicating their mechanistic role [35-37]. Furthermore, the efficacy of B-cell targeted therapies in RA through the depletion of B-cells, and hence autoantibody producing plasma cells, adds further credibility to the importance of autoantibodies [38].

Although animal models do not fully replicate human disease, they contribute to evidence supporting the pathogenic role of ACPA. Murine studies have demonstrated the induction of collagen induced arthritis and detection of ACPA following inoculation with citrullinated peptides [39-41]. The interaction of HLA shared epitope motif and ACPA has also been demonstrated to produce arthritis in murine models [42, 43]. Whether murine ACPAs have a role in either inducing or aggravating disease is debatable [44]. *In vitro* studies have demonstrated that human ACPA can activate the complement system [45]. In addition, interaction between macrophages and ACPA containing immune complexes resulting in the production of tumour necrosis factor (TNF) has been shown. This provides a potential mechanism for APCA driving inflammatory pathology [46].

More recently the role of ACPA in bone loss has been investigated. One group has reported that ACPAs have the potential to specifically bind to and activate osteoclasts [47]. The authors demonstrated the activation of osteoclasts and induction of bone loss following administration of ACPA isolated from individuals with RA. Furthermore, the binding of ACPA to osteoclasts and their precursors in murine models resulted in pain-like behaviours prior to the onset of inflammation. This offers a potential explanation for the presence of pain and arthralgia in at-risk individuals prior to the active inflammation of RA [48]. Precisely how the effect on osteoclasts contributes to inflammation in the synovium remains to be elucidated although several theories have been proposed including the release of cytokine IL-8 [49]. It is probable that whilst ACPAs are important, several other mechanism of immune stimuli are responsible for the establishment of synovial inflammation.

2.1.2.2 Genetic risk factors

As discussed earlier, for ACPAs to have a pathogenic role in disease initiation requires an individual to have the appropriate genetic background and exposure to environmental factors. Twin studies have been pivotal in providing evidence for genetic associations in RA. Data suggests the concordance for RA in monozygotic twins is between 15-30% and dizygotic twins 5%, with estimates for the heritability at around 60% [50, 51]. Subsequent studies suggest a slightly lower contribution, with first-degree relatives of patients with RA having a threefold increased risk for the development of RA [52, 53].

Large genome wide association studies (GWAS) have advanced our understanding of the genetic susceptibility to RA [54, 55]. A 2013 GWAS reported 101 genetic loci associated with the disease [56]. The majority of genetic associations relate to ACPA positive rather than ACPA negative disease. The most important genetic risk factor for RA remains the human leucocyte antigens alleles (HLA DRB1 *0101, *0102, *0401, *0404, *0405, *0408, *1001 and *1402) that encode a conserved amino acid motif known as the shared epitope (SE). Overall, HLA-DRB1 accounts for 40% of the genetic component of susceptibility of RA. The HLA DRB1 *04 allele has been shown to confer the greatest risk of RA development and increased further by the presence of two alleles [57].

Large association studies reviewing the single nucleotide polymorphisms (SNPs) across the major histocompatibility complex have identified the strongest signal to be at amino acid position 11, followed by position 71 and 74 [58]. Two other SNPs outside the HLA-DRB1 region have also been identified as important; amino acid position 9 of HLA-B in class I and amino acid position 9 at HLA-DPB1 in class II. These amino acids are located within the peptide binding cleft of the antigen presenting molecules. SE position surrounding and within the binding cleft provides a mechanistic role which is pivotal to the ‘shared epitope hypothesis’ [59]. Due to the SE linkage, there is successful antigen presentation, T-cell receptor activation and resultant immune response. Subsequent research has identified avid binding of citrulline-containing peptides compared to their unmodified counterparts in the binding cleft [43]. This accounts for SE association with ACPA positive disease. More recently, researchers have suggested that the SE may act

as a ligand favouring polarisation of T-cell differentiation towards an autoimmune T-cell repertoire – specifically the enhancement of Th17 differentiation and inhibition of a generation of T regulatory cells [60].

Studies have shown that HLA-DRB1 SE alleles not only contribute to disease risk but are implicated in severity [61]. Similar to disease susceptibility, individuals homozygous for HLA DRB1*0404 were four times more likely to have high disease burden with an erosive phenotype than SE negative individuals [62]. Not all the HLA alleles have been linked to disease risk, HLA DRB1*1301 has been demonstrated to be protective in ACPA positive RA [63].

There are genetic associations that have been found outside the HLA region. The most reported is protein tyrosine phosphatase non-receptor 22 (PTPN22) [64, 65], which is associated with increased risk of ACPA positive RA [66]. This gene encodes for lymphoid-specific protein tyrosine phosphatase which down regulates signal transduction from T-cell receptors. The variant risk allele (PTPN22 R620W - in which a single base change in the coding region results in an amino acid substitution) has been implicated in a number of other autoimmune disease [67, 68]. The PTPN22 polymorphism appears to be relevant to particular geographical or ethical groups and is absent from many Asian populations. Where present, it can relate to a two-fold risk of developing RA.

Although focus has been predominantly on ACPA positive disease, specific loci have been identified in ACPA negative disease including IRF5 (haplotype in the promoter region of the interferon regulatory factor 5 gene) and HLA DR3 [69, 70]. From GWAS it would appear that ACPA positive and negative disease display significant risk allele frequency differences within the HLA region, where as there is more overlap with non MHC associations [71].

The fact that concordance between monozygotic twins is, at most, 30% implies that other factors besides genetic predisposition act as contributors. Epigenetic and stochastic factors contribute in varying degrees to RA pathogenesis. A recent analysis

using a twin database [72, 73] allowed researchers to conclude that the development of ACPA relied largely on environmental and stochastic factors rather than the presence of HLA DR-SEs [74]. The authors suggest that the significant effect of HLA-DRB1 alleles is not at initial induction of autoimmunity, but in determining disease initiation, as illustrated by Figure 2. Several of these factors are now discussed.

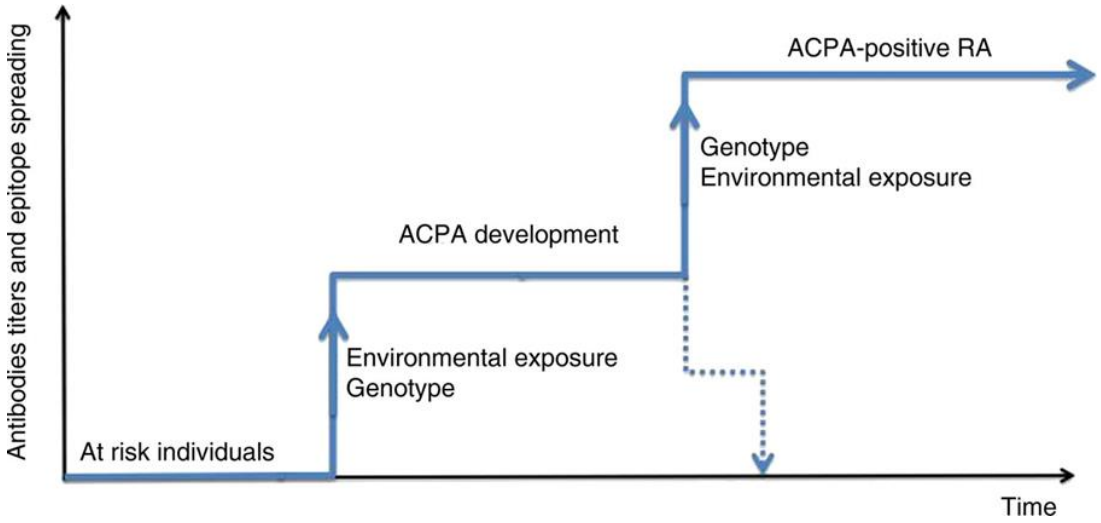


Figure 2 Schematic representation of a model for stepwise development of ACPA and ACPA-positive RA

Reprinted by permission from BMJ publishing group [74] Copyright 2015

2.1.2.3 Epigenetics

In addition to changes in genetic sequence, epigenetics or 'outside conventional genetics' refers to the study of potentially heritable alterations in gene expression that arise during development and cell proliferation [75]. Importantly with epigenetic processes there is no alteration to the underlying DNA sequence. Simply, it is any biological mechanism that switches genes on and off. Epigenetic processes include DNA methylation and posttranslational histone modifications (methylation, acetylation, phosphorylation, ubiquitinylation, SUMOylation and poly-ADP-ribosylation) which result in added residues or side chains marking regions of genetic code [76, 77]. These epigenetic markers are recognised by 'reader' proteins that serve as platforms and docking sites for effector proteins which facilitate processes such as transcription, DNA replication, damage response and chromatin remodeling.

The role of epigenetics in RA remains to be elucidated. It is acknowledged that epigenetics has an impact on the expression of cytokines, chemokines and metalloproteinases involved in inflammatory disease [78]. Initial epigenetic studies have focused on the key effector cells in RA, such as the fibroblast-like synoviocytes (FLS). RA FLS have been shown to have increased levels of histone deacetylase which has been shown to regulate the expression of cytokines such as TNF and IL-1 as well as account for the high cell proliferation and decreased cell apoptosis [79]. Researchers have demonstrated a therapeutic effect of histone deacetylase inhibitors on some models of arthritis. Whilst DNA methylation has been shown to repress transcription by inhibiting the binding of transcription factors and is disrupted to varying degrees in RA samples [80, 81].

Changes regulated by microRNAs (miRNAs) are another important area of epigenetics. MiRNAs are a highly-conserved class of short non-coding RNAs (21-25 nucleotides) that serve as transcriptional negative regulators involved in fine tuning of the expression of genes for cell differentiation, metabolism and immunity [82, 83]. It is now thought that they regulate at least 30% of messenger RNA [84, 85]. Several miRNA have been implicated in RA (reviewed [86]), with miR-155 and -146 being the most widely reported [87-89].

2.1.2.4 Environmental contributors

Potential environmental exposures, can be encompassed in the category of stochastic factors (random life events), which are thought to play an important role in the aetio-pathogenesis of RA. The contribution of environmental factors to health and disease has been studied in many autoimmune diseases [90]. Large observational studies have retrospectively identified several environmental risks for RA (Box 1) [91].

Box 1 Environmental factors that have been associated to risk of RA *[91-112]*

Factors generally thought to increase risk of RA
Silica exposure
Low UVB exposure

Smoking*
Low alcohol intake
High BMI‡
High birthweight
Lower socioeconomic status
Factors generally thought to decrease risk of RA
Longer breast-feeding duration
Factors with an indeterminate effect on risk of RA
Periodontal disease
Hormones—oral contraceptive§
Dietary components—consumption of red meat, proteins, oily fish, fruit, caffeine
Psychological stress
Low levels of vitamin D

**Particularly in seropositive disease.‡ particularly in seronegative disease. §Studies suggest a protective effect.*

Hormones

The predominance of autoimmune disease in females suggests that hormones and the reproductive system play an important role on disease development [113]. As suggested in Box 1, although initial studies suggested a protective effect with oral contraceptives, subsequent evidence has been inconclusive [96, 114, 115]. Recent, epidemiological studies have demonstrated that longer exposure to the pro-inflammatory sex hormones (early menarche) and pregnancy are associated with increased risk of disease development [97]. Pregnancy can induce remission in a high proportion (around 90%) of women, however relapse is frequent post-delivery [116, 117]. Furthermore, once disease has initiated, subsequent changes to hormone levels appears to effect disease severity with multiparty being associated with poorer outcomes.

Smoking

Cigarette smoking is the most widely accepted environmental risk factor for RA and demonstrates a significant interaction between environment and genetic factors [118, 119]. As with HLA-SE and PTPN22 risk alleles, smoking confers the greatest risk to ACPA

positive RA. Studies have demonstrated a multiplicative effect of genetic risk (HLA-SE) and smoking [119, 120]. Individuals who possessed two copies of the HLA-SE and smoked had a 21-fold increased risk of RA development compared to HLA-SE negative non-smokers [119].

The association of disease with smoking as well as the lung abnormalities documented in early RA has contributed to the theory that the lung mucosa may be an initiating site of disease [121, 122]. This hypothesis suggests that by-products of smoking contribute to the citrullination of peptides and hence generation of antigens which evoke an autoantibody response. The identification of inducible bronchus-associated lymphoid tissue (iBALT) in patients with pulmonary complications of RA supports the concept of localised autoimmunity [123]. Studies have demonstrated citrullinated antigens at the lung mucosa of healthy smokers compared to non-smokers [124]. More recently, mass spectrometry-based proteomic techniques have identified shared citrullinated antigens in both synovial and bronchial samples of RA patients [125]. Whilst histology from the lungs of untreated early RA individuals (with no apparent lung disease) confirmed lymphocyte aggregates with increased activation markers particularly in seropositive individuals [126]. Although a modest study the authors were able to demonstrate these changes were largely independent of smoking status suggesting increased citrullination at the lung is not solely attributed to smoking.

Involvement of the lungs prior to RA and systemic autoimmunity can be assessed by examining first degree relatives (FDRs) – these individuals share both genetic and environmental associations. Preliminary results investigating induced sputum, as a marker of lung autoimmunity, demonstrated a greater incidence of autoantibodies in the sputa compared to sera [127]. The role of the lung mucosa in RA has incited interest and forth coming studies will hopefully provide clarification.

Periodontal disease

An emerging environmental risk factor for the development of RA is periodontal disease and the plethora of microbes found in the mouth [128]. The discovery that one of the

main organisms associated with periodontal disease, *Porphyromonas gingivalis*, citrullinates peptides via peptidyl arginine deiminase (PAD) enzymes, thereby forming the antigens required for ACPA formation, suggested an association between periodontal disease and RA [129]. However, population-based studies report conflicting results for this disease association, with confounding factors including smoking and poor oral hygiene as well as discrepancies in quantifying periodontal disease by self-reported methods [108, 130-132]. Investigative studies have attempted to determine whether the presence of *P gingivalis* precedes RA onset or is merely a bystander. Thus far, evidence suggests titres of antibodies to *P* gingivalis are associated with the presence of RA-related autoantibodies in at-risk individuals [133-135]. It can be hypothesized that the generation of ACPA via PAD enzymes contributes to the antibody response and triggers systemic disease in a proportion of individuals.

Gut Microbiome

As detailed in Box 1, specific dietary associations with RA remain inconclusive but research has examined the role of the gastrointestinal system microbiome [136, 137]. The microbiome refers to microorganisms which populate the human body, primarily relating to the skin and mucosal organs. The interaction between the immune system and infectious agents is not a new theory, there has previously been work evidencing the link between bacteria and viruses including the Epstein Barr virus, proteus species and Escherichia coli [138-141]. Whilst the exact mechanism has not been proven, molecular mimicry is postulated [142].

Within the gut there are several notable observations from murine models, the first reported in the 1970s. This study described how arthritis prone rats raised in a germ free surroundings developed severe inflammatory arthritis whilst those in conventional cages (with environmental pathogens) developed a milder form of disease [143]. Subsequent murine studies have attributed protective and pro-arthritic effects of pathogens [144-146]. More recently, in germ free IL-1 receptor antagonist-knockout mice, researchers demonstrated that only those colonized with defined microbiota (*lactobacillus bifudis*) developed rapid onset severe arthritis similar to that seen in non-germ free rat models [147]. This was shown to be due to an imbalance in T-regulatory

cells and Th17 cells. Investigation into T-cell repertoires within these studies suggests that gut bacteria maintain an immune-homeostasis between pro- and anti-inflammatory T-cell subsets. Dysbiosis with overgrowth of pathogenic bacteria is hypothesised to disrupt the homeostasis; thus contributing to RA development [148].

In humans, bacterial flora in individuals with early RA appears to be distinct to that found in health, other diseases and untreated RA patients, with an abundance of *Prevotella copri* with loss of Bacteroides [149, 150]. Whilst no pathogen has been identified capable of citrullination, the notion that dysbiosis and a triggering of gut immune cells may have a pathogenic role, which has led to increased research in the microbiome and autoimmunity. Large scale international projects are underway to attempt to map out the human microbiome and provide potential answers to its role in health and disease [151, 152].

The concept of environmental factors triggering autoimmunity and an inflammatory response in susceptible individuals has been widely accepted (Figure 3). However, the precise cause of the loss of immune tolerance and the perpetuation of this state remains unclear. It may be that particular microbial products, microvascular change, biomechanical stress, or most probably a combination of factors, contributes to the initiation of synovitis [23].

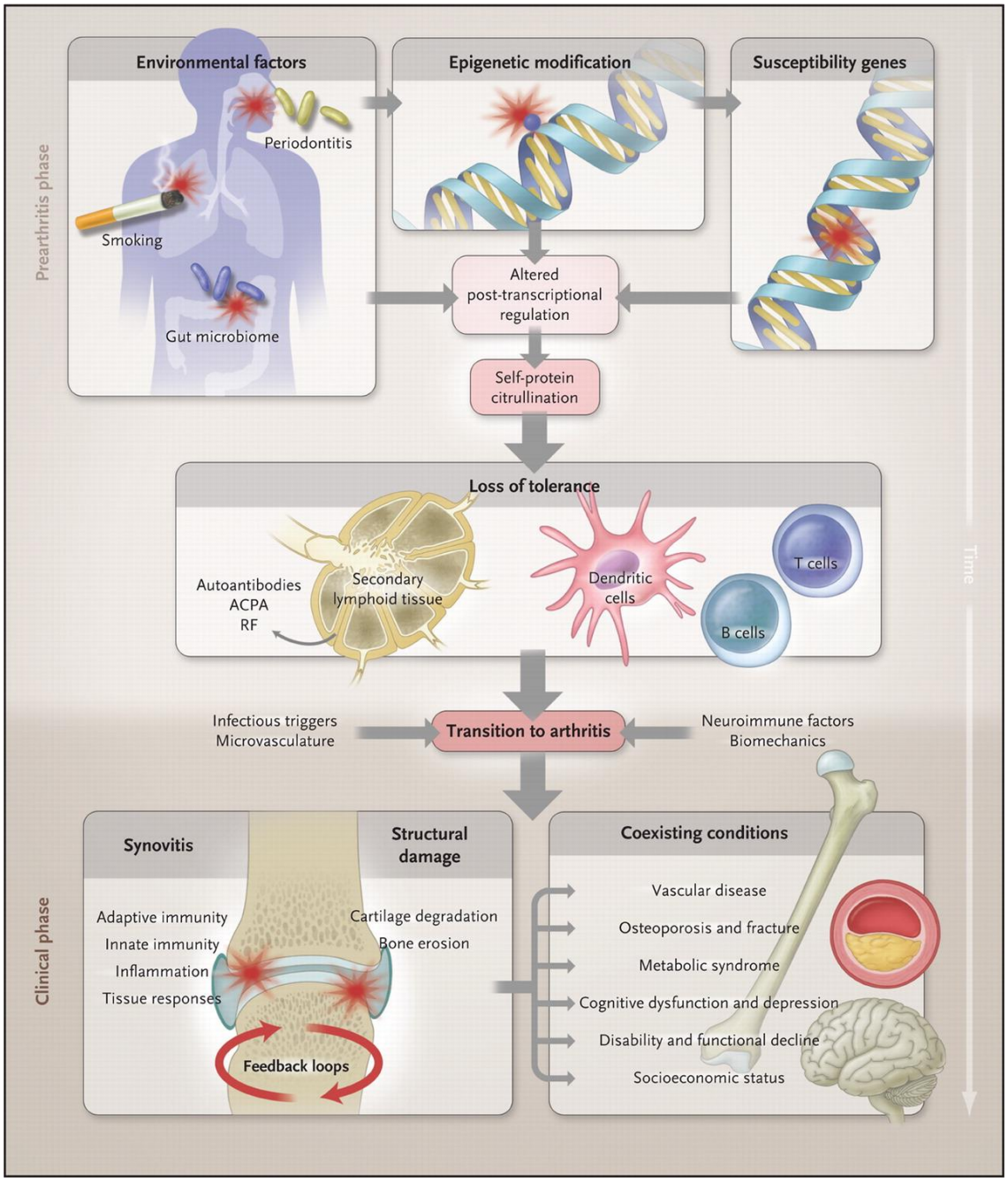


Figure 3 Multistep process and potential triggers to inflammatory arthritis development.

Reproduced with permission from [23], Copyright Massachusetts Medical Society.

2.1.3 The inflammatory arthritis continuum

The natural history of RA can be considered along a continuum with distinct stages of progression (Figure 4). Here, an individual with the genetic predisposition to disease can be followed through to exposure of environmental factors and the onset of immune mediated disease – an inflammatory arthritis such as RA.

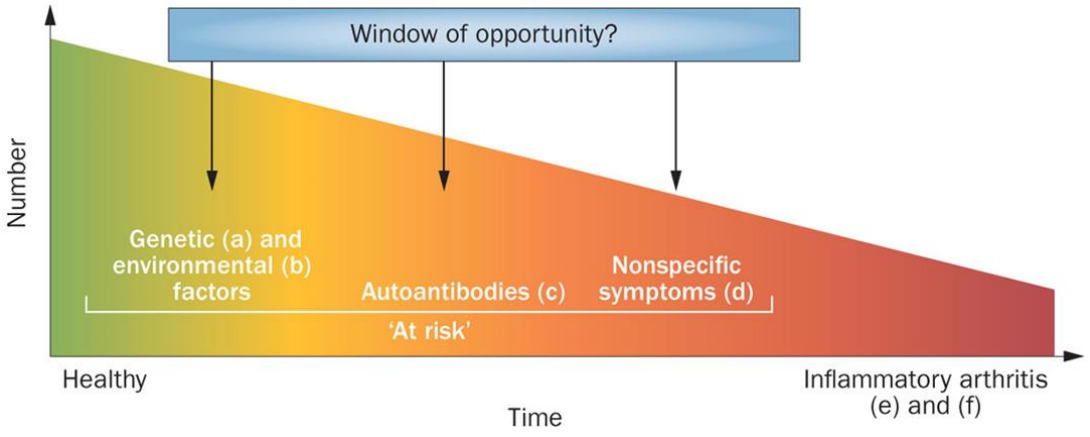


Figure 4 Schematic of inflammatory arthritis continuum

Reproduced from review article [7].

As illustrated in the inflammatory arthritis continuum there are several stages prior to the diagnosis of RA (Figure 4 a-d). The figure references the categorises recently suggested in the recommendations for terminology of 'at-risk' groups [153]. Individuals may possess one or several factors that are classified as a risk. Genetic (a) and environmental (b) interactions may interact to increase an individual's risk of developing IA. An asymptomatic period marked by the presence of systemic autoantibodies (c) and other inflammatory markers may occur, enabling the early identification of 'at-risk' individuals. Once clinical disease is established individuals can be broadly defined as undifferentiated arthritis (e) and RA (f).

At the point of clinically detectable disease, the disease course can be heterogeneous namely; persistent, intermittent, relapsing, progressing or remitting. Often the ability to give an exact diagnosis is not possible. van Aken *et al* suggested that up to a third of individuals with a new IA will be classified as undifferentiated (UA) [154]. One definition for UA is 'an early form of arthritis not meeting classification/diagnostic criteria for a more definitive disease' [155]. From large inception cohorts of IA, the outcome of individuals with UA largely reports around one third will develop RA while in up to a half the IA will resolve completely [156, 157].

2.1.4 Diagnosis

The diagnosis of RA is based on a clinical assessment involving clinical history, examination findings and exclusion of alternative differentials diagnoses. Laboratory tests can also help the diagnosis process and have been included in both the ACR 1987 and EULAR/ACR 2010 classification criteria for RA [158, 159]. The new criteria in 2010 (Box 2) was developed following reports of a relatively low sensitivity for the 1987 criteria detecting early disease [160]. These new criteria have already been evaluated in several cohorts. A recent systemic review pooled this data producing an overall sensitivity 82% and a slightly lower specificity 61% [161].

Box 2 The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis

Target population (Who should be tested?): Patients who….

1. have at least 1 joint with definite clinical synovitis (swelling)
2. with the synovitis not better explained by another disease

Classification criteria for RA (score-based algorithm: add score of categories A–D; a score of ≥6/10 is needed for classification of a patient as having definite RA)

	SCORE
A. Joint involvement	
1 large joint	0
2-10 large joints	1
1-3 small joints (with or without involvement of large joints)	2
4-10 small joints (with or without involvement of large joints)	3
>10 joints (at least 1 small joint)	5
B. Serology (at least 1 test result is needed for classification)	
Negative RF *and* negative ACPA	0
Low-positive RF *or* low-positive ACPA	2
High-positive RF *or* high-positive ACPA	3
C. Acute-phase reactants (at least 1 test result is needed for classification)	
Normal CRP *and* normal ESR	0
Abnormal CRP *or* abnormal ESR	1
D. Duration of symptoms	
<6 weeks	0
>6 weeks	1

Adapted from [158].

With regards to laboratory markers with positive serology for autoantibodies, RF and in particular ACPA, are central to the diagnosis of RA. Whereas RF can be found in other diseases and healthy individuals, ACPA has a high specificity for RA and is present in 60-80% of individuals with RA [162]. Raised inflammatory markers/acute phase reactants such as C - reactive protein (CRP) and erythrocyte sedimentation rate (ESR) are typically seen in new presentations of RA often in conjunction with thrombocytosis and leucocytosis. An abnormal acute phase reactant gains one score in the new classification criteria (Box 2).

The steering group responsible for the new criteria for RA also provided guidance with regards to bone erosions. This is particularly important given the widespread use of radiography and ultrasonography in early arthritis clinics. Conventional radiographic examinations remain the 'gold standard' for diagnosing joint damage in RA. Whereas, previously erosive disease was one of the criteria for a diagnosis of RA which resulted in delays in diagnosis, the new criteria do not stipulate the need for erosion. In the presence of clinical synovitis, ultrasonography can be used to confirm and detect subclinical synovial involvement in other joints which may help direct treatment.

Applying the new classification criteria to historical cohorts and registries has demonstrated that individuals who were previously diagnosed as UA at presentation would now be classified as RA. The Norfolk Arthritis Registry has reported higher incidence rates of RA in both males and females following application of the new classification criteria in the inflammatory polyarthritis cohort [163]. Although, the overall incidence of RA has been reportedly decreasing, it will be instructive to monitor these trends following adoption of the 2010 criteria and the change in definition of disease.

2.1.5 Prognosis

Current treatment strategies aim to treat disease aggressively to prevent joint deformity, disability and the psychological implications of chronic disease [164-166].

However, RA is a heterogeneous condition and can entail a variable course. One treatment that is effective for one individual will not necessarily be effective for another. Identification of prognostic factors associated with poor outcome in disease progression and persistence is a priority area of clinical research. Of equal importance is the identification of individuals with a less active disease phenotype in whom milder treatment may induce remission with a tailored management pathway. Whilst this concept of personalised medicine has not been fully adopted in early arthritis clinics, research has identified presenting features which are associated with protracted disease and poorer outcomes.

Several studies have investigated prognostic factors in RA from early arthritis cohorts. Clinical factors which have been attributed to a poor prognosis include; erosive disease [167], high number of active joints and subcutaneous nodules [168]. Gender appears to be inconclusive, although it is generally accepted that females have poorer outcomes [169, 170]. Smoking has also been shown to correlate with erosive disease and poorer outcomes [171, 172] which, as previously discussed is hypothesised to be due to the interactions with ACPA [119]. ACPA titre has been cited to be the greatest contributor in a prediction model for disease severity. This reported a ten times increased likelihood of erosive disease compared to ACPA negative patients [173]. Other laboratory measurements including high levels of inflammation markers (CRP and ESR) are associated with poorer prognosis [174]. Male gender, the absence of erosions and good functional scores (e.g. HAQ) have been cited as good prognostic markers for remission, while the absence of RF and low number of active joints at diagnosis equate to more favourable outcomes [167, 175, 176].

Patients with undifferentiated arthritis are at risk of development of a persistent IA fulfilling the criteria for RA. Researchers have attributed several prognostic markers for development of RA, by virtue some of these are those associated with poor RA outcomes. A prediction rule for progression to RA from UA at one year was developed using data from an early arthritis inception cohort [156]. Nine factors were identified as predictors; age, female gender, distribution of affected joints (favouring classical bilateral poly-arthropathy phenotype), tender and swollen joint counts, duration of

morning stiffness, level of CRP, and RF and ACPA positivity, these have been validated in other cohorts [155].

Although a relatively new area of research, significant works evaluating the stages prior to overt inflammatory disease (UA and RA), the 'preclinical stage', have been reported. This has been driven by the concept of aiming to treat disease at the earliest opportunity to obtain best outcome. The concept of preventing arthritis is now being considered [177, 178]. However, before prevention can be implemented an understanding of the stages prior to disease onset and the risk of progression to RA is required.

2.2 Classification of 'at-risk' cohorts

The study group for risk factors for RA have defined the key terminology for the different phases of disease (Box 3) [153]. An individual may fulfil one or more of the classifications e.g. have both genetic and environmental risk factors. Using this resource it is apparent that there are several defined cohorts that could be considered for prevention. This classification provides a suitable foundation for defining 'at-risk' cohorts and individuals that may be amenable to preventative strategies. Importantly, that the term pre- RA can only be used retrospectively once an individual is known to have developed RA.

Box 3 Terminology recommended for the phases prior to RA development [153]

- In prospective studies individuals would be described as having:
  - (a) Genetic risk factors for RA
  - (b) Environmental risk factors for RA
  - (c) Systemic autoimmunity associated with RA
  - (d) Symptoms without clinical arthritis
  - (e) Unclassified arthritis
  - (f) RA
- The term 'arthritis' is used to denote clinically apparent soft tissue swelling or fluid (not bony overgrowth alone).
- (a) to (e) can be used in a combinatorial manner for example, an individual may have (a)+(b), or (a)+(b)+(c) or (a)+(b)+(d), etc.

▸ The prefix 'pre-RA with:' can be used before any/any combination of (a) to (e) but only to describe retrospectively a phase an individual was in once it is known that they have developed RA.

The at-risk category originally considered were those individuals identified due to the presence of systemic autoimmunity through analysis of biobanks of stored sera (at-risk category (c)). First degree relatives fall into several categorises of risk ((a)+(b)+(c)) and are easily identifiable through RA probands. By virtue of increased presentation to physicians; individuals with arthralgia or symptoms of arthritis and autoimmunity are a further category that has also been evaluated ((c)+(d)). Further discussions in this chapter will focus on these categories of risk, focusing specifically on the latter ((c)+(d)).

2.2.1 Individuals with systemic autoimmunity

Predictors identified in nested case control cohorts

Several clinical and serological characteristics are associated with an increased risk for the development of RA. Over 20 years ago, del Puente *et al* reported on the presence of autoantibodies prior to disease. By prospectively following an asymptomatic cohort in a population known to have a high risk of RA, rheumatoid factor (RF) was demonstrated to be a predictive factor for future development of RA [179]. However, it has been the availability of stored blood samples permitting retrospective testing for the presence of autoantibodies (reviewed in [180]), which has enabled the greatest insight. Initially studies focused on the presence of RF isotypes but later included other markers such as antifillagrin and antikeratin- now known to be part of the antibody response to citrullinated peptides [181-186].

The importance of these autoantibodies was highlighted when Swedish and Dutch blood biobanks reported on the positive predictive value (PPV) [34, 187]. By adopting a nested case-control method, they analysed serial blood samples in individuals prior to clinical disease. The presence of RF isotypes and citrullinated cyclic peptides antibody (anti-CCP) were compared to that of age-matched controls. Prior to clinical diagnosis, 40.5%

of the Dutch and 33.7% in the Swedish study group were anti-CCP positive. The presence of autoantibodies predated the onset of RA by up to 14 years. The 5-year PPV for anti-CCP in the blood donor population was 96.6% and 100% if IgM RF and anti-CCP were present. Translating this to the general population, a positive anti-CCP had a 5-year PPV for the development of RA of 5.3%. The frequency of positive samples increased closer to the onset of clinical disease in both cohorts, supporting the concept of a mounting immune response. This is consistent with studies that have reported an expanding repertoire of citrullinated antigens evoking autoantibodies in the years prior to disease [35, 36, 188, 189]. Whether specific antigens are directly related to development of RA has yet to be clarified.

Several large studies in America have adopted the case-control design to identify risk factors for chronic disease, primarily cancer and cardiac disease, but also RA. The Nurses' Health Study is one of the largest having had two periods of recruitment 1976 and 1989 with respective recruitment and prospective follow up of 121,700 and 116,430 women [190]. Although the main objective was to consider lifestyle choices and environmental factors contributing to disease, over a quarter of participants in each cohort provided a blood sample. As with the European biobanks, these studies have evaluated many biochemical, immunological and environmental factors that may be associated with RA and have demonstrated a mounting immune response from cytokines, autoantibodies and inflammatory proteins prior to RA development [184, 188, 190-198]. Owing to the wealth of data that is available, researchers can adapt their study to consider potential new biomarkers or risk factor as the field evolves. One recent study from the Netherlands demonstrated increased bone markers in the serial samples obtained prior to the diagnosis of arthritis. This suggested that there is an alteration in bone metabolism, with increased osteoclast activity during the systemic autoimmunity phase [199]. Whilst, the Nurses Health Studies' team analysed vitamin D intake and levels (a topical and debated risk factor of autoimmune disease) [102, 200]. No significant relationship was demonstrated, although there appeared to be a trend towards low levels of vitamin D in the months prior to RA diagnosis in a subset of individuals. These and other similar studies have contributed to our understanding of

the immune, metabolic and inflammatory response prior to disease although there are limitations given reliance on medical records and self-reported questionnaires.

Predictors identified in first degree relatives:

Similar to findings originally reported with the Pima Indians, RF presence was shown to predate diagnosis of RA in FDRs [201]. These individuals often share genetic and environmental factors with their RA relatives. As expected, this group of individuals had a higher prevalence of ACPA and RF isotypes compared to healthy controls [37, 202-207]. From the blood donor studies, individuals considered as 'high risk' (2 family members with RA) had a PPV for the development of RA over 5 years of 69.4% if anti-CCP was detected, increasing to 100% if IgM RF was also present [208]. The isotype profiles appeared to be different to RA populations, with several reports highlighting a predominance of IgA ACPAs [204, 205]. The significance of this has yet to be determined but may indicate the mucosa as a site of pathogenesis (discussed Section 2.1.2).

A large cohort of RA probands has been recruited in the United States as part of the Studies of the Etiology of Rheumatoid Arthritis project [207]. This cohort has enabled researchers to study known as well as possible novel associations and risk factors for the development of RA. Similar to the retrospective blood donor studies, these databases are now reporting on other predictors such as cytokine patterns in FDRs [209, 210]. More recently they reviewed dietary supplementation of omega 3 fatty acids and reported a protective effect for the presence of ACPA positivity in probands [211]. A similar cohort is now being recruited in the UK [212]. In other autoimmune diseases, FDRs have been the focus of preventative trials [213-215]. This is not currently proposed in RA, although it is possible that these individuals would be amenable to health promotion intervention to avoid the known risk factors of RA.

2.2.2 Individuals with systemic autoimmunity associated with RA and symptoms without clinical arthritis:

Individuals with musculoskeletal symptoms and autoimmunity (Box 3 (c) and (d)) provide a key cohort for understanding of inflammatory arthritis. Importantly these individuals can be assessed clinically and followed prospectively for progression to RA. To date, there are two key cohorts which have reported on progression rates and these individuals' characteristics [216, 217]. The initial prospective study detailed the characteristics of 147 participants who tested positive for IgM RF and/or ACPA in the presence of arthralgia symptoms [216]. A 20% (29/147) rate of progression to arthritis in a median of 11 months was reported, with 34.5% fulfilling 1987 ACR criteria for RA [159]. Of those who progressed, 90% (26/29) were ACPA positive, giving ACPA positivity a hazard ratio of 6.0. IgM RF was only related to progression in the presence of ACPA. The Leeds cohort recruited individuals with non-specific joint pain and ACPA positivity resulting in an even higher progression rate of 50% [217]. As the cohorts and duration of observation has increased, several characteristics have been evaluated and predictors of progression identified.

'Biomarkers' is a term frequently used in health sciences when commentating on characteristics of disease or outcome. A definition for a biomarker has been provided;

> *'A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention*.' [218]

A biomarker in the field of prediction can therefore be a physical sign, cellular, genetic, or biochemical/molecular alteration that aids diagnosis or prognosis. Additionally, for a biomarker to have clinical utility it should be easily obtainable, highly reproducible, sensitive, specific for the outcome and where possible reflect a pathogenic process [219]. These specifications are important to consider when translating findings from at-risk cohorts in to day to day clinical practice.

2.2.3 Biomarkers described in at-risk cohorts

Table 1 illustrates several biomarkers that have been identified for evaluation of at-risk individuals with systemic autoimmunity and symptoms of arthritis.

Biomarkers	Number of cases	Summary of findings
Serological immune markers		
▪ ACPA & IgM RF [216]	147	90% of those who developed arthritis were ACPA positive (HR 6.0, 95% CI 1.8–19.8). In ACPA-positive individuals, the addition of RF but not SE enhanced the risk of arthritis progression. No independent association between arthritis progression and RF positivity or presence of SE.
▪Fine Specificity of ACPA [220] *(Reactivity against 5 different peptides derived from fibrinogen, enolase and vimentin analysed)*	244	82% of anti-CCP positive individuals and 9% of anti-CCP-negative, RF-positive individuals recognised >1 citrullinated peptide; these patients had an increased risk of developing arthritis compared with patients who recognised 0–1 peptides (OR 2.1, 95% CI 1.0–4.4). A specific citrullinated peptide associated with progression to arthritis was not identified.
▪ anti-CarP antibodies [221]	340	Of those positive for anti-CarP antibodies (36% of the study population), 51% developed RA compared with 25% of those who were negative. Presence of anti-CarP antibodies is associated with an increased risk of arthritis progression independent of anti-CCP and RF status (HR 1.56, CI 95% 1.06–2.29)
▪ FC glycosylation to ACPA [222]	183[§]	Using serial samples from an ACPA+ arthralgia cohort demonstrated a decrease of galactosylation and an increase of core fucosylation of serum ACPA-IgG1 shortly before the onset of RA.
Gene expression & inflammatory markers		

▪ Gene expression [223]	109	Signatures associated with arthritis development were involved in IFN-mediated immunity and chemokine/cytokine activity (OR 21.0, 95% CI 2.8–156.1). Genes involved in B-cell immunology were associated with protection against arthritis progression (OR 0.38, 95% CI 0.21–0.70)
▪ Interferon signature [224]	115	IFN signature scores generated by measuring expression of 7 type I IFN response genes demonstrated IFN(high) signature was associated with progression to IA in cox regression analysis HR 2.38 (95% CI 1.26 to 4.49; p=0.008) after correction ACPA and RF status.
▪ B cell signature [225]	115	Expression of CD19, CD20, CD79α and CD79β measured to create a B cell signature. B cell (high) revealed protection compared to B cell (low) signature (HR 0.50, 95% CI 0.26 to 0.98, p=0.042). Significance lost after correction for RF and ACPA status.
▪ HsCRP, PCT, SPLA2, TNF-alpha, IL-6, IL-12p70, IL-10, and IFN- and several mRNA biomarkers [226]	137	A trend was observed with some cytokines but, overall, circulating cytokines did not differ between those who progressed to arthritis and those who did not. No correlation was found between mRNA expression levels of inflammatory genes and progression to arthritis
Lipid profiles		
▪ TC, HDLc, LDLc, TG, apo A1 and apoB [227]	348	Differences in lipid profiles were noted between those who developed arthritis and those who did not, although the difference was not significant for all markers. A decrease in apoA1 was predictive of development of arthritis in anti CCP positive individuals only (HR 0.52, 95% CI 0.29–0.92).

▪ Adipokines in serum and synovium [228]	51	Increased serum Vaspin levels was associated with an increase in progression to IA (HR1.5 95% CI 1.1 to 2.2; p = 0.020) studied in 27 autoantibody positive individuals. Synovial expression of adipokines was not associated with progression to IA.
Histology markers		
▪ Tissue architecture: Synovial knee biopsy [229]	13	Explorative study – No difference in tissue architecture detected between individuals who progressed to arthritis and those who did not.
▪ Synovial tissue expression [230]	15	Synovial expression of Prostaglandin E_2 pathway enzymes were not associated with arthralgia symptoms or progression to IA.
▪ Lymph node phenotype [231]	20	Lymphoid pro-inflammatory CD8+ T-cells exhibit a less-responsive phenotype in at-risk individuals. Increase in CD8+ memory T-cells in LN accompanied by an increase in non-circulating or recently activated (CD69+) CD8+ T-cells in LN and matched peripheral blood compared to health. No risk to progression analysis as no development to IA reported.
▪ Lymph node phenotype [232]	12	LN assessments in at-risk and RA groups. Innate lymphoid cell (ILC) profile in the LN changes from a homeostatic towards a more inflammatory profile during the at-risk and earliest phase of RA. No risk to progression analysis as no development to IA reported.
Imaging markers		
▪ MRI of knee [229]	13	Explorative study – No difference in inflammation detected on imaging between individuals who progressed to arthritis and those who did not.

▪MRI of hands [233]	21[§]	Subclinical inflammation detected on MRI (particularly in the wrist) in ACPA-positive patients with arthralgia, but studies with larger numbers of patients and prospective follow-up are required to determine whether this finding predicts development of arthritis.
▪MRI of hands and feet [234]	28[§]	MRI synovitis was present in 93% ≥1 joint(s) and was not associated with progression to IA, median [IQR] cumulative synovitis score in those that progressed 9.0 [6.3-13.8] compared to 20 [10.8-21.8] in those that did not, OR 0.87 (95% CI 0.76-0.99). There was a temporal relationship between extent of synovitis and time to IA, with those with a score 2 of synovitis progressing sooner than a score 1.
▪MRI of hands [235]	20[§]	Synovitis, osteitis and bone erosions were demonstrated in 65%, 35%, and 65% respectively. Tenosynovitis was the most reported pathology present in 80% but none of the controls. Individuals with ≥2 tendons involved were more likely to develop RA.
▪US of hands [236]	192	There is a trend of US abnormalities (joint effusion, synovitis, PD and tenosynovitis) in those who progressed to arthritis. This did not reach statistical significance. Results suggest US better at detecting subclinical inflammation compared to clinical examination.
▪US of hands [237]	136[§]	PD was predictive of progression to IA; PD score 2 in a joint equated to a 30-fold increase in risk of developing arthritis at that joint (HR 31.3, 95% CI 15.6, 62.9), p<0.001. At a patient level, progression also occurred sooner with a PD score of 2 compared to a zero; median time to clinical arthritis 7.1 months verses 52.4 months, HR 3.7 (95% CI 2.0,6.9) p=<0.001. Individuals with erosions also progressed sooner.

▪PET of hands [238]	29[§]	Explorative study – 4 patients had at least 1 PET positive joint. All these individuals progressed to arthritis. Of the remaining 5 who progressed but had negative PET scans 3/5 developed inflammatory arthritis in joint not included in the PET scan. Suggests PET therefore has a clinical utility at detecting subclinical arthritis.
▪Micro-CT [239]	15[§]	Bone mineral density measured by micro-CT was significantly reduced in ACPA positive individuals compared to healthy controls. Analysis of risk of progression not performed due to small number.
Association with environmental factors/clinical factors		
▪Smoking and BMI [240]	55	Independently, smoking (HR 9.6, 95% CI 1.3–73.0) and high BMI (HR 5.6, 95% CI 1.3–25.0) were associated with arthritis development. Individuals with both a BMI ≥25 kg/m2 and a history of smoking had a higher risk of developing of arthritis compared to never smokers with a BMI ≤25 kg/m2. The increased risk associated with smoking was found to be independent of ACPA status.
▪Alcohol intake [241]	361	Alcohol consumption at baseline was inversely related with the risk of development of arthritis; HR 0.80 95% CI (0.65 to 0.97). No associations identified concerning the quantity of alcohol consumed.

Table 1 Biomarkers considered in individuals with systemic autoimmunity (anti-CCP and or RF) and arthralgia

[§] All ACPA+ individuals +/-RF (remainder of studies consisted of individuals IgM RF+ or ACPA+)

Anti-CCP: anti-citrullinated cyclic peptide antibody, ACPA: anti-citrullinated protein antibodies, Anti CarP: Anti-carbamylated protein antibodies , apo A1/B apolipoprotein A1/B, BMI: body mass index, HDLc: high-density lipoprotein cholesterol, HsCRP: High-sensitivity CRP, IL-6: interleukin 6, ILC: innate lymphoid cell, INF-: interferon-, LN: Lymph node, LDLc: low-density lipoprotein cholesterol, MRI: magnetic resonance imaging, mRNA: messenger RNA, PCT: procalcitonin, PD: power

Doppler, PET: Positron Emission Tomography, RF: rheumatoid factor, SE: Shared epitope, SPLA2: secretory phospholipase A2, TG: triglycerides, TNF-alpha: tumor necrosis factor-alpha, TC: Total cholesterol, US: ultrasound

2.2.3.1 Serological immune markers:

Blood sampling is a minimally invasive method of identifying potential characteristics which may in the future be used as biomarkers. Autoantibodies have been one of the first biomarkers considered in the at-risk cohorts [35, 216, 220]. More recently novel autoantibodies to carbamylated protein (anti-CarP) have been reported in both RA and at-risk cohorts [221, 242, 243]. Studies have reported the role anti-CarP in diagnosis and prediction, particularly in ACPA negative RA. Anti-CarP presence demonstrated an independent association to RA progression with an at-risk cohort. Anti-CarP has been assayed in the Leeds ACPA+ at-risk cohort as part of a collaborative project [244]. Of the 123 ACPA+ at-risk individuals tested, 40.6% (50/123) were positive for anti-CarP; 58% (29/50) of these progressed to RA during follow-up. A weak association with disease progression was demonstrated HR=1.70 (95% CI 0.91, 3.18), p=0.099. Although anti-CarP has been demonstrated to be present before RA in the biobank samples [245], further validation in prospective at-risk cohorts is required to ascertain its utility in the preclinical setting.

A potential immune marker yet to be considered in at-risk populations are the circulating T-cell and B-cell populations. The role of the immune system - in particular lymphocytes, in RA pathogenesis is supported by several factors: i) the presentation of antigens by MHC to immune cells, ii) autoantibody and cytokine production, and iii) infiltration of lymphocytes in the synovium [59, 246]. Whilst T-cell subsets have been considered in RA for prediction remission, relapse and response to therapy [247-250] they have not been considered in at-risk cohorts and offer a potential new predictor of disease.

2.2.3.2 Genetic markers:

Genetic markers have also been considered in at-risk cohorts. One group has reported that the HLA-shared epitope (SE) allele is not independently associated with disease progression [216]. However, specific gene signatures, particularly those related to interferon mediation, have been identified [224].

There have however been no reports concerning microRNA (miRNA), a novel biomarker that has been evaluated in cancer extensively. MiRNAs are short non-coding RNAs (21-25 nucleotides) which regulate hundreds of mRNAs and are one of the key epigenetic factors in disease pathogenesis (Section 2.1.2.3). This regulatory characteristic combined with their relative stability and presence in various biological samples renders them highly suitable as potential biomarkers. Specific miRNA related to inflammation in RA have been reported, although functional work is still insufficient [251, 252]. Other pre-disease states such as Barrett's oesophagus progressing to oesophageal adenocarcinoma, illustrate the potential for specific miRNA signatures to be associated with disease risk [253].

2.2.3.3 Histological markers:

Histological samples provide insights into disease pathogenesis. Obtaining such samples can be invasive and potentially unacceptable to the general population outside of research settings. However, attitudes towards biopsies would change if the clinical value of histological markers were appropriately highlighted to patient groups and clinicians. The benefit of synovial samples has been illustrated in early IA cohorts in which disease stratification was possible [254-256]. Additionally, synovial pathotype has assisted in the development of peripheral biomarkers which predicted response to biologic therapy [257]. The clinical utility in at-risk cohorts would be considered high, if they can inform risk and assist in tailoring management.

Within the at-risk cohorts, researchers have studied synovial and lymph node tissue [229-232]. Studies evaluating the synovium have demonstrated infiltration of T-cells [229]. In at-risk cohorts synovial studies have been limited largely due to the lack of clinically swollen joints accessible for biopsy. Furthermore, synovial tissue sampling is a highly technical procedure that requires experienced personnel to ensure appropriate tissue is selected with minimal discomfort to the participant [258]. Reassuringly, a recent review of synovial biopsy procedures reported an overall complications rate of 0.4% [259]. Ethical considerations are pertinent in these cohorts particularly if the joint is not inflamed nor symptomatic.

Lymph node biopsies have also been attained from at-risk individuals. In this study, the newly recognised cell line believed to mediate immunity – the innate lymphoid cells (ILC), was reported to be present [260, 261]. The authors demonstrated significant differences between the ILC subsets between health, at-risk and RA disease groups [232]. However, numbers recruited were small and no progression data was available within the at-risk group for the authors to comment on progression risk. The insights into disease pathogenies these studies provide remain greatly valued.

2.2.3.4 Imaging markers:

Imaging modalities have enhanced our understanding of inflammatory disease. Clinical imaging has good biomarker potential given that several imaging modalities now provide dynamic and macroscopic mapping of tissue without the inconvenience and possible harm of performing biopsies.

i) Radiographs

Conventional radiographs are frequently used to assess for the presence and monitoring of peri-articular damage and erosions in RA. In early RA it is possible to detect changes in joint damage within 3 months with plain radiographs [9]. Radiographic changes at baseline are predictive of disease progression in undifferentiated cohorts and in individuals with RA predicts future radiographic damage [262, 263]. However, given the subtlety of changes that occur prior to disease it is not a sensitive modality to assess changes within the joint. This is likely to explain why there is little data currently available regarding radiographs of individuals with autoimmunity but no clinical synovitis.

ii) MSK Ultrasonography

Although initially reserved to research use, ultrasound of the joints is now well-established and common place in many rheumatology departments and early arthritis

clinics and is recommended in the assessment of rheumatic diseases [264, 265]. Studies have demonstrated that ultrasound is more sensitive than clinical examination [266, 267]. Visualising the synovium and its vascularity has aided early diagnosis and assessment of disease activity. Ultrasound is also more sensitive at detecting bone erosions compared to conventional radiographs. The ability to examine a joint in multiple planes, and the function of power Doppler (PD) assist in the diagnosis of early inflammatory disease.

Within at-risk individuals, there have been two key studies which have reported ultrasound findings [237, 268]. A study of 192 individuals (ACPA positive and/or RF positive arthralgia) demonstrated ultrasound abnormalities (joint effusion, synovitis and PD) to be predictive of development to clinical synovitis at a joint level (OR 3.07 (95% CI 1.05, 8.94), OR 5.45 (95% CI 2.32, 12.8), and OR 5.50 (95% CI 2.32-12.8) respectively [268]. Combining the presence of synovitis (grade 2-3) and PD (grade 1-3) increased the risk of development to arthritis in that joint to OR 12.9 (95% CI 4.65-36.0) translating to a positive predictive value of 35%. However, ultrasound was not found to be predictive of progression to arthritis at a patient level, although there was a positive trend. In contrast, Nam et al reported ultrasound to be predictive at both a patient and a joint level [237]. In this study, a core set of 32 joints were scanned in 136 anti-CCP-positive individuals. The ultrasound results were kept blinded from the clinical assessors to reduce any bias in assessment of clinical synovitis. Thirty percent had PD in ≥1 joint(s), 96% had GS in ≥1 joint(s) 21% had ≥1 erosion(s). Interestingly, there was a high prevalence of grey scale change found in the MTPs of healthy controls (n=48) and was shown to be less discriminating in the anti-CCP-positive cohort between those who progressed and those who did not. Therefore, the MTP scores were excluded. A baseline scan with a PD score of 2 in a joint equated to a 30-fold increase in risk of developing arthritis at that joint (HR 31.3, 95% CI 15.6, 62.9), $p<0.001$. At a patient level, progression also occurred sooner with a PD score of 2 compared to a zero; median time to clinical arthritis 7.1 months verses 52.4 months, HR 3.7 (95% CI 2.0,6.9) p=<0.001. Similarly, individuals with an erosion in at least one joint were at greater risk than those without; median time to clinical arthritis 7.5 months verses 50.1 months, HR 2.9 (95% CI 1.7,5.1), p=<0.001.

iii) Magnetic resonance imaging (MRI)

MRI has advantages over other imaging modalities and this is reflected in its recommended use in diagnosing RA in recent EULAR recommendations [269]. MRI can visualise the field of interest in 3 orthogonal planes providing detailed assessment of bone and surrounding soft tissue whilst avoids ionising radiation. Therefore, it follows that investigators considered MRI for the at-risk groups to ascertain any subclinical disease.

One of the first studies to report MRI findings was conducted in the Netherlands and matched synovial biopsy of the knee to MRI findings of the same joint [229]. This small explorative study (n=13) of individuals with autoantibodies but no clinical synovitis (12/13 having arthralgia) demonstrated no difference in MRI measure compared to the healthy control group (n=6). The negative findings could reflect the joint imaged given that knee is rarely the first joint involved in RA, furthermore just under half subjects reported symptoms at the joint. However, the authors defend their conclusions extrapolating from data from synovial samples from unaffected RA joints, in which increased synovial inflammation is found [270]. However, in disease initiation, imaging (and biopsy) of the small joints might provide more insight.

Imaging of the small joints suggests early changes can be seen on MRI in at-risk individuals, albeit, thus far, in small cohorts [233, 234]. Krabben *et al* demonstrated that individuals with autoantibodies and arthralgia had higher mean inflammation scores at the wrist joints (sum of synovitis and bone marrow oedema scores using OMERACT and RAMRIS system) compared to healthy controls; 0.9 vs 2.3 respectively ($p<0.001$ *confidence intervals not reported*) [233]. Interestingly, in the 21 arthralgia autoantibody positive group the mean inflammation score in the painful joints was 1.0 and in the symptom free 1.2 suggesting symptoms did not reflect MRI features. Furthermore, during follow up 12 individuals developed RA with the inflammation scores not differing from those who did not progress to RA.

Similarly, a study analysing 28 autoantibody positive arthralgia patient's MRI scans concluded the presence of MRI synovitis was not associated with progression to clinical arthritis [234]. In fact, those individuals that progressed to disease state had synovitis scores lower than those who did not. However, there did appear to be a substantive relationship when the authors considered time to synovitis in those with a greater degree of synovitis in one joint. Individuals with synovitis score of 2 in at least 1 joint developed arthritis within 1 year, and those with scores of 1 developed disease more gradually. These studies demonstrate that MRI is sensitive at detecting subclinical inflammation but the specificity remains questionable when considering its role in predicting progression to arthritis.

iv) *Positron emission tomography (PET)*

PET is a modality infrequently used in clinical practice when assessing individuals with RA. Small studies have suggested PET imaging offers greater sensitivity at subclinical synovitis compared to MRI [271] and potentially detect those at risk of progression to RA [272].

v) *Microcomputer tomography (micro CT)*

Similarly, micro CT is a modality currently reserved for research purposes to assess bone microstructure and density. Since the presence of ACPAs and RF have been demonstrated to equate to a greater risk of bone erosions in RA [273-276], researchers have attempted to assess bone loss in healthy individuals with detectable ACPA compared to controls [239]. Here, the ACPA positive group (n=15) had significantly greater reduction in bone mineral density and cortical bone thickness with distinct changes to cortical bone architecture. This programme of work challenges the notion that synovial inflammation is seen prior to the activation of osteoclasts and bone resorption. Instead, the concept of ACPAs possessing a pathogenic role in which they contribute to a reduction in bone integrity before the milieu of cytokines and inflammation become established is suggested. Although yet to be demonstrated in vivo; in vitro, ACPAs from RA patients were shown to bind to osteoclasts with a resultant

induction in osteoclastogenesis and bone-resorptive activity [47]. Whilst a multitude of questions still remain, at-risk individuals provide a unique opportunity to study concepts regarding the pathogenesis of RA.

2.2.3.5 Associations with environmental factors:

Smoking has been identified as a risk for progression to RA in 55 individuals with arthralgia and systemic autoimmunity [240]. This study also considered obesity which has previously been reported to correlate with an increased risk of future diagnosis of RA in the Nurses' Health Study [277]. In this prospectively followed cohort, both 'ever' smoking and body mass index (BMI) ≥25 kg/m^2 were associated with arthritis development; HR (95% CI): 9.6 (1.3 to 73.0); p=0.029 and HR (95% CI): 5.6 (1.3 to 25.0); p=0.023 respectively [240]. There were no differences in the titre or antibody status between individuals in the two groups. This analysis included only 55 individuals, and validation of these associations in larger cohorts is warranted.

The role of periodontal disease, in particular *P gingivalis,* has been studied within these individuals. To date, assessments of the immune response in the sera to *P gingivalis* have failed to show a risk to disease progression but has established a significant association with the presence of RA-related autoantibodies in individuals at-risk [135, 278]. This may be due to insufficient powering of studies and the methods used to assess for periodontal disease and *P gingivalis*. Recently data from population based case-control studies in the Swedish biobanks have reported increased levels *P gingivalis* related antibodies in the serial samples prior to disease [279, 280]. Here, the antibody response to the *P gingivalis* virulence factor arginine gingipain were measured as opposed to merely *P gingivalis* antibodies. Given that chronic periodontitis is estimated to affect approximately 30% of the population, measuring antibodies to *P gingivalis* did not necessarily provide a mechanistic link [281]. Ginigpains are necessary for the cleaving of peptides and hence generation of *de novo* epitopes for citrullination by PPAD [282]. Establishing a causal link between periodontal disease, the oral microbiome and initiation of disease remains challenging but several studies combing clinical and microbial sampling are hoping to provide further evidence [283].

Alcohol has been reported to be inversely associated with the development of RA [112, 284] and likewise within at-risk cohorts autoantibody positive individuals with no alcohol intake had a greater risk to RA development than drinkers [241]. However, within our own cohort we did not see a protective effect [217].

From the biomarkers identified in Table 1, it is evident that it may now be possible to identify individuals at potential risk of progression to arthritis in the systemic autoimmunity and symptoms risk category. However, to date, results are limited to a few cohorts based in the northern hemisphere. Data from other geographical regions and ethnicities will provide further information and are warranted.

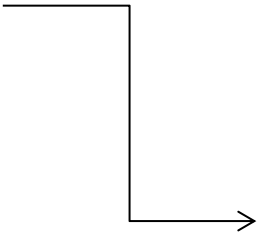
2.3 Risk stratification and model development:

As documented, no single biomarker has been identified which predicts disease with sufficient accuracy to be of clinical value. A different approach is therefore needed. The ability to risk stratify individuals is an attractive option particularly in light of current strategies concerning personalised medicine [285]. This has been undertaken in the early phases of inflammatory disease where models have been developed which aid prediction of progression from UA to RA [156, 286]. However, recent trials which aimed to prevent progression from UA to RA did not risk stratified individuals and may account for failure to reach primary endpoints. A recent re-evaluation of the 'PROMPT' study in which individuals were treated with methotrexate versus placebo demonstrated that by risk stratification of patients, those at high risk would have hypothetically resulted in better outcomes [287]. This illustrates the importance of risk stratification prior to offering treatment to individuals at risk of RA.

Risk score tools have been developed by determining clinical risk factors associated with RA in an autoantibody positive (RF and/or ACPA) cohort. One recently published scoring system, reported a score greater than 7 corresponded to a risk of developing RA within 1 year of 43% and within 5 years 81% [288]. This scoring system used a minimum of 9

variables, which is a relatively large number of data points for performing risk stratification. Some of the variables used may only be recorded in the local research clinic in which the model was designed. The transferability to other cohorts for validation has therefore been limited. Risk stratification of ACPA positive individuals at the Leeds research clinic, resulted in a simple tool to be used in primary care and rheumatologists to identify 'at-risk' individuals early [217]. Several clinical, serological and imaging markers were considered in the first 100 patients. Variables found to be substantively associated with progression in univariable analysis were included in the multivariable models of progression. Ultrasound imaging along with selected clinical findings and antibody titres resulted in a model in which no individual determined as low risk progressed to IA, compared to 72% if categorised as high risks score. Figure 5 illustrates the scoring system.

Criteria	Points
Tenderness of small joints present	1
EMS ≥ 30 minutes	1
High level RF and/or anti-CCP	1
Power Doppler present	1
Shared epitope present	1



Total Score	Proportion progressing of those who progressed to IA within 12m	
0	0/3 (0%)	Low 0%
1	0/8 (0%)	
2	7/25 (28%)	Mod 50%
3	9/31 (29%)	
4	10/19 (53%)	High 72%
5	4/6 (67%)	

Figure 5 Risk stratification model including clinical, serological and imaging biomarkers. Demonstration of categorisation of individuals using the clinical risk score.

This model facilitates the identification of individuals at greatest risk, in whom the use of immune modulating therapies could be targeted to prevent disease progression. The results from this analysis require validation in a larger patient cohort.

2.4 Intervention studies in 'at-risk' cohorts:

Until very recently, the closest this area of research had to an interventional study were large, population-based studies in which the effect of various (non-DMARD) therapeutic options were analysed to assess for any reduction in RA development [289-292]. The Women's Health Study prospectively assessed the effect of low-dose aspirin and vitamin E supplementation in the prevention of CVD and cancer. A sub-analysis of this same cohort demonstrated that both interventions did not reduce the incidence of RA [289, 292]. Similarly, a randomised control trial evaluating the effect of postmenopausal hormone therapy on CVD, hip fracture and breast cancer outcomes, performed an analysis of the incidence of RA [290]. This study, which included 27,347 participants, reported a non-significant reduction in the risk of developing RA (HR 0.74; 95% CI 0.57–1.10) with the use of postmenopausal hormone therapy. These large population-based studies provided information on how prevention therapies could be targeted. Other dietary and supplementary therapies have been assessed in similar cohorts, and could provide insights on how prevention strategies for RA could be developed.

2.4.1 Corticosteroids

As understanding of the preclinical phase advanced, investigators attempted to delay and prevent progression to RA. Bos *et al* recruited individuals with arthralgia and autoantibodies (either ACPA and/or IgM RF) and treated with intramuscular glucocorticoids at 0 and 6 weeks [293]. The endpoint chosen was 50% reduction or normalisation in autoantibody levels at 6 months. This was achieved in one patient from each group. Although this intervention was hoped to have an impact on the development to RA, long-term follow-up revealed similar percentages in each group.

2.4.2 DMARDs

'Strategy to Prevent the Onset of Clinically-Apparent Rheumatoid Arthritis' (StopRA) is a randomised placebo control trial which has recently began recruitment in the United States [294]. Individuals with positive anti-CCP titre are recruited and randomised to receive hydroxychloroquine or placebo for 12 months. The end point is the

development of rheumatoid arthritis as defined by the ACR/EULAR criteria (score≥6) at 36 month follow up. Individuals will be recruited from health fairs, FDRs of RA and rheumatology clinics. No evidence or history of an IA or prior DMARD use are essential to the exclusion criteria.

2.4.3 Biologic DMARDs

Recent reports from trials in early IA and RA advocate the use of early biologic agents with induction of remission in 69% to 89% [295, 296]. The ability to maintain remission on a reduction or even withdrawal of biologic drug suggests the immunological effect can be sustained if disease is treated early. It can be hypothesised that treating prior to disease onset, those at high risk may have a sizeable immunological impact. There are currently two trials underway exploring prevention in antibody positive individuals; the PRAIRI study (NTR 1969) using a single dose of 1000mg of rituximab and APPIPRA study (EudraCT Number: 2013-003413-18) 52 week therapy with abatacept. Each of these studies have been designed acknowledging current understanding regarding the drugs mechanisms of action and evidence from RA or UA trials. The use of rituximab, a depleting anti-CD20 therapy, is effective and well tolerated in early and late stages of RA and shows particular efficacy against autoantibody positive disease [297, 298]. B-cells role in the production of immunoglobulins including RF and ACPA, present many years before the onset of clinical disease, is one of the reasons for its application in this preclinical phase. The presence of B-cells and plasma cells in the synovium of active RA patients also supports its potential role in disease pathogenesis. Furthermore, B-cells can produce cytokines and can trigger autoimmunity due to their antigen presenting cell (APC) function and hence activation of T-cells through co-stimulatory signals. Similarly, abatacept is commonly used therapy in the treatment of both early and late stage RA [299, 300]. This fusion protein (consisting of Fc region of immunoglobulin IgG1 and CTLA-4 molecule) binds to CD80/ CD86 molecules on the T-cell surface preventing activation. The hypothesis is that preventing APCs from delivering the co-stimulatory signal at this pivotal stage of disease initiation could prevent progression to the full phenotype of clinical RA.

2.5 Summary

In this chapter, the potential initiating factors involved in the pathogenesis of RA are reviewed. The concept of a continuum with distinct stages to the natural progression of disease is detailed. Evidence indicating that patient outcomes are improved through early intervention has fuelled research into early disease identification. Subsequently, the opportunity to prevent disease has become an attractive option. Determining onset of disease and optimal intervention time point has now become a research priority. In order to address this need, the stages prior to disease and individuals at risk to RA are pivotal. Those with systemic autoimmunity (ACPA+) offer insights into disease pathogenesis. The fact that not all individuals progress to arthritis exemplifies the heterogeneity within this at-risk cohort. Hence, researchers have attempted to identify predictors of disease. This chapter has reviewed the serological, imaging and clinical factors which have been considered. As yet, no single biomarker has been identified as solely predictive. However, the potential to risk stratify individuals using multiple biomarkers is a viable option. This approach would enable the identification of those at greatest risk and thus, in whom exposure to therapeutic agents are justifiable. Ongoing research investigating disease pathogenesis and risk identification enables the opportunity for innovative therapeutic approaches, strategies for health promotion and ultimately concepts in disease prevention to be explored.

2.6 Hypothesis and Aims

The overarching hypothesis of this thesis states that the development of rheumatoid/inflammatory arthritis can be predicted in at-risk individuals be evaluating imaging, molecular and cellular biomarkers. Subsequently, individuals at risk to IA/RA can be risk stratified to assist clinical management.

By studying individuals presenting with systemic autoimmunity and non-specific musculoskeletal pain this work aims to:

I. Evaluate the role of MR imaging in the prediction of IA/RA development.

II. Explore the change in MiRNA expression along the IA continuum and the biomarker potential in predicting IA/RA development.

III. Establish whether T- cell subset dysregulation is associated with progression to IA/RA and assists risk stratification models.

3. Methods: Study Design and Population

The study design and population are detailed in this chapter. Clinical and imaging assessments are described. Specific methods and statistical analysis relating to each work stream (e.g. laboratory methods of cellular markers) are detailed in the relevant results chapter.

3.1 Ethical Approval

The study was approved by the Leeds Ethics Committee (research ethics committee reference: 06/Q1205/169). Approval was granted in November 2006. There have been subsequent amendments, the most recent of which in 2014.

3.2 Study Design

The *'Co-ordinated Programme to Prevent Arthritis: Can We Identify Arthritis at a Pre-Clinical Stage?'* is a longitudinal prospective cohort study. The study comprises of two components. The primary care component is adopted by the UK National Institute of Health Research Clinical Research Network. Recruitment was initially limited to Yorkshire, however subsequent approval has been granted at recruitment sites throughout the UK.

General practitioners, musculoskeletal physicians, physiotherapists, podiatrists and other health professionals are asked to refer individuals presenting with new, non-specific musculoskeletal (MSK) joint pain. This includes presentations such as lateral epicondylitis, bursitis and shoulder tendonitis. The primary care component comprises of the completion of health questionnaires and blood testing for anti-CCP2. Those participants who are anti-CCP positive are eligible for the second component of the study, forming the 'at-risk' cohort.

3.3 Consent

Participants are provided with information and given a minimum of 24 hours before informed consent is obtained. They are able to discuss the study with the study team by contacting the research helpline telephone number. Informed consent process is confirmed by the signing of the consent forms by the patient and a good clinical practice trained research nurse or doctor. Participants recruited to the main study are invited to take part in the affiliated sub-studies.

3.4 Study Population

3.4.1 At-risk cohort

In this programme of work, the at-risk cohort (or CCP+) are defined as those individuals with non-specific MSK joint pains, who have tested positive for the anti-CCP2 assay. In addition to those identified by primary care, eligible individuals were referred from the early arthritis clinics within Yorkshire. Box 4 outlines the study inclusion and exclusion criteria. Patients included in these analyses were recruited from June 2009 to January 2016.

Box 4 Study inclusion and exclusion criteria.

Primary & Secondary Care subjects
Inclusion Criteria • Age > 18 years • Has a new musculoskeletal complaint • Capable of understanding and signing an informed consent form • Has tested CCP positive
Exclusion Criteria • Patient fulfils 1987 ACR Criteria or the 2010 ACR/EULAR criteria for RA • Has a definitive diagnosis of an inflammatory arthritis on referral • Patient has tested CCP negative

3.4.2 Control Subjects:

Healthy controls were recruited to the study. Baseline assessments were performed following the same protocol as the subjects with non-specific MSK joint pain and autoimmunity (CCP+). Healthy control subjects attend for one baseline assessment.

3.5 Study Protocol

The study protocol comprises of a main clinical and imaging study and a biological blood sample sub-study (stored serum). Table 2 details the research protocol with reference to study visits, clinical parameters, imaging and sampling. In addition to the listed time points, a participant could attend outside of these scheduled visits should they develop new signs or symptoms consistent with inflammatory arthritis.

Table 2 Participant Schedule listing study requirements for each visit.

Study Visit **Week**	**1** **0**	**2** **13**	**3** **26**	**4** **39**	**5¥** **52**	**6** **Withdrawal**
Inclusion / Exclusion & Consent	X					
Medical, family & social history	X					
Examination, observations - height, weight, BP, pulse	X					
Complaints-directed examination		X	X	X	X	X
Adverse Events/Medications	X	X	X	X	X	X
Joint count	X	X	X	X	X	X
Early morning stiffness	X	X	X	X	X	X
Patient Questionnaire : HAQ, VAS, Employment Questionnaire, EQ5d	X	X	X	X	X	X
ANA, RF	X				X	X
Anti-CCP test	X	X	X	X	X	X
ESR, CRP, Haematology	X	X	X	X	X	X

Chemistry, HLA DR4	X					
Stored serum	X	X	X	X	X	X
Urine	X				X	X
X-rays[§], MRI[β]	X				X	x
High Resolution US	X		X		X	X

§ X-ray of hands and feet and symptomatic joints as clinically indicated at baseline and repeat at 12 months to a maximum of 4 joint regions per visit.

β In patients with no contraindication to imaging and as facilities allow. Failure to undergo these investigations does not exclude participant from remainder of study. MRI to be performed at time points above unless individual participating in another investigational study in which an MRI is to be performed.

¥ The first year of this longitudinal study is shown in the above schedule. This is an on-going data collection study and subsequent visits after this year will continue at annual intervals with the same format as visit 5 (excluding MRI). If clinical need dictates, the above procedures may be performed at more or less frequent intervals.

3.6 Assessments

3.6.1 Clinical & demographic assessments

Clinical assessments were performed at 3 monthly time points for the first year and then annually or as clinically indicated. Initial baseline visit allowed collection of demographic and medical history data. Early morning joint stiffness (EMS) and current symptomatology (e.g. intermittent symptoms, tendonitis, and arthralgia) were recorded at each study visit. Participants completed patient questionnaire during visits including health assessment questionnaire (HAQ), visual analogue score (VAS), employment questionnaire and a measure of health-related quality of life measure (EQ5d).

Rheumatologists experienced in the identification of IA (LH, CR and KM) performed the examinations. Assessments included a musculoskeletal examination, 78 joint count for tenderness and 44 joint count assessing for swelling. Physicians were also able to comment and record other findings such as tendonitis. The rheumatologists were blinded to the results of the imaging investigations (MRI and ultrasound).

Patients were provided with the research helpline number to enable contact with the research team should new symptoms of IA develop. This allowed rapid review by the research team rheumatologists and appropriate study investigations to be instigated.

3.6.2 Imaging Assessments

3.6.2.1 Conventional radiography

X-rays of hands and feet were performed at baseline. Follow-up x-rays were performed at 1 year and then as clinically indicated. Any symptomatic joints may be x-rayed if clinically appropriate. Musculoskeletal radiologist reported on the presence of abnormalities.

3.6.2.2 High Resolution Ultrasound Scan Protocol

Examinations were performed by a rheumatologist and an ultrasonographer trained in musculoskeletal ultrasonography on a Philips HDI 5000 machine 5–12 and 8–15 MHz transducers (later changed to a General Electric S7 machine). Power Doppler was assessed using a pulse repetition frequency set between 700 and 1000 MHz. The rheumatologist and ultrasonographer performing the scans were blinded to the clinical examinations.

Standard protocol required a total of 38 joints to be scanned which consisted of the wrists, metacarpophalangeal joints (MCPs), proximal interphalangeal joints (PIPs) and metatarsophalangeal joints (MTPs) bilaterally (Appendix A). Following a preliminary analysis of MRI data which indicated that tendon pathology may be an important finding, tendon scoring of the hand flexor tendons (at the MCP and PIP level) and the extensor carpi ulnaris was introduced. The OMERACT definitions were used to define synovitis [301]. For scoring EULAR OMERACT system was applied (semi quantitative scale) [302].

In addition, if patients had any symptomatic joints that were accessible by ultrasound these could be scanned and appropriately scored. Following baseline assessment, joints were scanned at 6 months and then annually or at the development of IA.

3.6.2.3 Magnetic Resonance Imaging Protocol

MRI scanning was performed on a Siemens MAGNETOM Verio 3.0 Tesla whole body scanner (Siemens Healthcare: Erlangen, Germany). Contrast-enhanced MRI was obtained after intravenous injection of 10 mL of gadolinium diethylene-triaminepentaacetic acid. Subjects were placed in the prone position with the hand extended in front of the body. MRI was performed at baseline and then, if possible, at the development of IA.

The protocol included coronal STIR of the hand and wrist to include the MCP joints with dedicated fat suppressed coronal T2 images separately acquired of the MCP joints and wrist, 3D gradient echo (Double Echo Steady State - DESS) with water excitation, T1 weighted 3D gradient echo (spoiled gradient echo) with water excitation pre and post gadolinium. The 3D sequences were acquired in the coronal plane with isotropic voxels allowing multiplanar reconstruction. Semi-quantitative scoring was performed independently by two musculoskeletal radiologists for synovitis, bone marrow oedema (BME) and erosions according to the OMERACT RAMRIS [303]. Tenosynovitis was scored using a previously described method [304, 305]. The 5 flexor tendons of the digits and the wrist flexor tendons were scored. For reference Figure 6 illustrates the tendons of the wrist. RAMRIS/OMERACT MRI score sheets can be found in Appendix B & C.

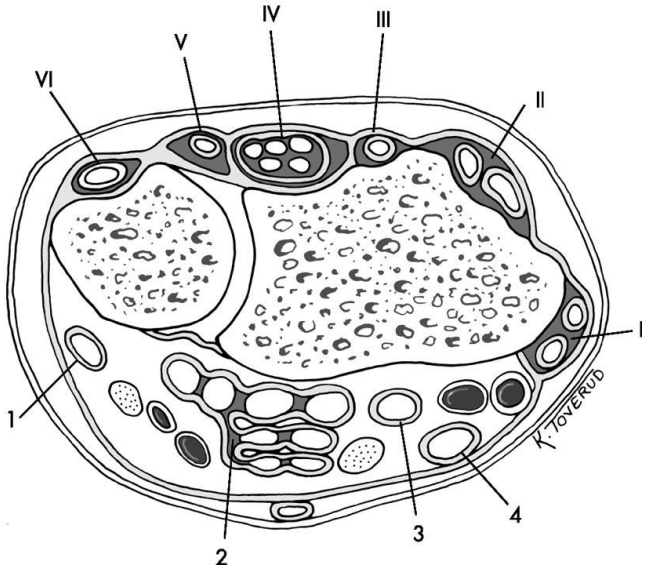


Figure 6 Illustration of the anatomy of the wrist extensor and flexor tendons**.**

Extensor compartments denoted in Roman numerals from I to VI: (I) extensor pollicis brevis, abductor pollicis longus; (II) extensor carpi radialis brevis, extensor carpi radialis longus; (III) extensor pollicis longus; (IV) extensor digitorum communis, extensor indicus proprius; (V) extensor digiti quinti proprius; (VI) extensor carpi ulnaris. Flexor tendon areas denoted in numbers from 1 to 4: (1) flexor carpi ulnaris; (2) ulnar bursa, including flexor digitorium superficialis and flexor digitorium profundus; (3) flexor pollicis longus (4) flexor carpi radialis.

Reprinted by permission from BMJ publishing group, [306] copyright 2007.

Patients were eligible for MR imaging permitting no contraindications had been identified and the safety questionnaire completed. The following exclusions applied; presence of pacemaker, surgical clips within the head, certain inner ear implants, neuro-electrical stimulators or metal fragments within the eye or head. Additionally, subjects were excluded if significant renal impairment as assessed by an estimated glomerular filtration rate <45ml/min/1.73m^2. Given the potential for allergic or anaphylactic reaction to gadolinium contrast, individuals with a history of multiple drug allergies or anaphylaxis were also excluded.

3.6.3 Blood sampling

Routine haematology and biochemistry were requested at specified time points of the study. This included full blood count, kidney/liver function tests, and inflammatory markers (CRP,ESR) which were reported using standard reference ranges as per local hospital trust. In addition, serology for RF (measured by nephelometry in IU/ml) and an

anti-nuclear antibody screen (multiplex bead technology, bioplex) giving titres of antibodies to clinically relevant antigen were collected and reported. A commercially available assay for anti-CCP2 test was used (initially immunocap 250, (Phadia) reference range <7U/mL) and later cohort re-evaluated using Bioplex 2200 machine (Bio-rad), reference range <2.99U/mL.

HLA DR typing to enable shared epitope status reporting was performed at baseline assessment by the Transplant and Immunology Services at St James University Hospital. SE (low resolution) was considered positive with the presence of one or two copies of the following alleles: HLA- DRB1*01, DRB1*04 and DRB1*10 in the HLA-DRB1 locus.

3.6.3.1 Biological Bloods Sub study

In the protocol schedule (Table 2) this item is listed as *stored serum*. A maximum 73mls of blood at baseline and 43mls at subsequent specified time points were collected. Blood was drawn into a combination of EDTA, sodium citrate, lithium heparin, serum clot activators, and PAXgene/RNA tempus tubes. After collection blood samples are processed for serum, plasma, or used fresh for flow cytometry and other functional studies, or stored (-20°C & -80°C) as whole blood for future DNA, RNA extraction depending on biomarker types and methods/techniques applicable.

The most recent study protocol amendment in 2014 allowed the annual blood draw from consenting participants following their initial 12 month follow-up.

3.7 Data Acquisition

Source data is recorded in patients' health records and study participant folders. Case report forms (CRF)/external electronic data is entered in a Microsoft share point page which has been set up for the study. Regular backups of the electronic data are performed.

Bloods tests performed under the NHS are available on a password protected results server. Research laboratories data is stored on the university server and contains no patient identification.

3.8 Data Storage

The electronic CRF is published and available for download as an excel on the NHS Trust protected system. No identifiable details are published.

3.9 Missing Data

Every effort will be made to clear missing data, including case note review. Multiple imputations will be employed in the event of any unrecoverable data.

4. Magnetic Resonance Imaging in individuals with systemic autoimmunity and arthralgia: MRI as an imaging biomarker

This chapter describes the use of MRI in an at-risk cohort to identify characteristic features. An evaluation of MRI as a predictive biomarker for disease progression is reviewed.

4.1 Introduction

MRI is a sensitive imaging technique that has advantages over both clinical examination and conventional radiographs for assessing joint damage (bone erosions, cartilage loss and joint space narrowing) and inflammation (synovitis, BME and tenosynovitis), which are common features in the earliest stages of RA [307-313].

When MRI has been used to assess small joints in early RA, approximately 70% of patients have bone damage at clinical presentation [314]. Tendons involvement has also been implicated in early disease with up to 75% of early RA individuals demonstrating MRI detected tenosynovitis [310, 315]. One of the advantages of MRI over other modalities is the ability to evaluate for the presence of BME (osteitis). This has been associated with progression and prediction of structural damage [266, 269, 316-318]. A randomised trial involving 256 methotrexate-naïve patients with RA, demonstrated that high baseline synovitis and osteitis were independent predictors of radiographic progression at 12 months [319]. Other studies also revealed an association between the severity of RAMRIS synovitis and BME with cartilage damage in RA [320, 321]. BME has been associated with other poor prognostic markers including ACPA and RF [322, 323].

MRI findings in at-risk populations:

As reviewed in section 2.2.3, MRI has been considered for differentiating those individuals at risk of RA development. However, studies have thus far, only investigated small cohorts (n≤28) which may account for the lack of consistency in the findings. The initial studies suggested that individuals with arthralgia and autoantibodies had more

subclinical MRI detected inflammation compared to healthy controls [233, 234]. Whilst the authors were unable to correlate baseline MRI features of inflammation to risk of progression, further analyses did suggest that the degree of subclinical inflammation was greatest in those at closest proximity to IA development [234].

As alluded to in these at-risk groups, there have been insufficient numbers for any robust analysis to ascertain whether MRI findings can predict progression to arthritis. This is primarily due to the difficulty in recruiting and imaging at-risk individuals. Using an alternative approach focused on signs and symptoms rather than autoantibody presence, van Steenbergen *et al*, devised a study in which individuals who were deemed to have clinically suspect arthralgia (CSA) were assessed [324]. Baseline imaging reported in 44% (41/93) of patients with CSA had MRI detected subclinical inflammation of joints of the hand and feet; these individuals would not have been identified through clinical and serological markers alone. Patients were categorised as MRI positive or negative dependent on the abnormalities observed in the healthy control reference group. By dichotomising the values for MRI inflammation to present or absent; 66 patients (45.8%) had a positive MRI. Univariate analysis of 142 patients demonstrated that higher MRI inflammation scores were associated with development of arthritis [325]. The presence of MRI inflammation at baseline was associated with progression to arthritis (HR 6.12, 95% CI 2.32, 16.19, $p<0.001$). This increased incrementally with the number of joints/bones involved (HR 1.23 per additional positive joint 95% CI 1.13, 1.33, $p<0.001$). All three MRI features (BME, synovitis and tenosynovitis) were significantly associated with arthritis development with tenosynovitis showing the greatest association (HR 7.56 95% CI 3.30, 17.32, $p <0.001$). Results of the multivariable analysis (Table 3) demonstrated MRI inflammation to be independently associated with arthritis development. Presented are the HRs of the multivariable analysis involving 142 patients with CSA that underwent MRI of which 27 (19%) developed IA. The sensitivity and specificity of MRI positivity within the CSA cohort was 81% and 63% respectively. Within the CSA population, individuals with no MRI inflammation had a low chance of developing arthritis (6%).

Table 3 Results of multivariable Cox regression analysis of clinical and serological factors and MRI-detected subclinical inflammation at baseline in relation to arthritis development.

	HR (95% CI)	p Value
Age, per year	0.96 (0.93 to 0.996)	0.028
Localisation of initial symptoms		
Small joints only	Ref	Ref
Large joints only	2.35 (0.41 to 13.61)	0.34
Small and large joints	4.30 (1.70 to 10.86)	0.002
CRP level, per mg/L	1.05 (1.01 to 1.09)	0.021
ACPA-positive	6.43 (2.57 to 16.05)	<0.001
Presence of any MRI-detected inflammation	5.07 (1.77 to 14.50)	0.002

The authors presented data from 124 individuals with CSA that underwent MRI of which 27 developed clinical arthritis. They sought to ascertain if the association of subclinical MRI inflammation with arthritis was independent of the association of the other listed factors. Results suggest an increased hazard for younger individuals, patient with initial localisation of symptoms in small and large joints compared to small joints only (reference), patient with higher CRP, ACPA positivity and patient with subclinical MRI inflammation. Ref, reference

Adapted from [325]

In this study, ACPA positivity was strongly associated with progression (HR 6.43, 95% CI 2.57, 16.05), emphasising the importance of this high risk group. Sixty three percent of ACPA positive individuals developed arthritis within a year (15/24). Within the ACPA positive subgroup, the chance of progression to arthritis, if subclinical MRI inflammation was present was 71% (PPV). However, in those with a negative MRI, 60% did not develop an arthritis within a year (NPV). Unfortunately, the small number of ACPA positive individuals with MRI findings limits the ability to comment on MRI performance compared to other markers.

MRI in other inflammatory arthritis populations:

Individuals with UA share clinical and imaging characteristics with the at-risk cohorts (inflammatory arthritis continuum Figure 2). MRI studies in these individuals aimed to

identity predictors of persistence or progressive disease. Findings from several studies have confirmed that MRI, including tenosynovitis, can predict progression from UA to RA [315, 326-329]. Authors have also sought to construct prediction models to assist physicians in evaluating and managing their patients [329, 330]. Similar principles apply in the at-risk populations and therefore form the basis for studies predicting progression to arthritis.

MRI finding in healthy individuals:

For MRI to be used as a diagnostic tool for early detection of RA, knowledge of the prevalence of MRI abnormalities in health is essential. It is recognised that MRI changes are more frequent in older patients [331]. However there were until recently, no published reference ranges for MRI abnormalities within the general population. The majority of data that informed a recent review came from MRI studies with control groups [332]. The reviewers summarised that erosions (RAMRIS ≥1) occur in 33-55% of healthy individuals (symptom free) dependent on joint(s) considered. Synovitis prevalence showed the greatest variation between the studies evaluated (0%-44%). Given the paucity of data in healthy cohorts, a study of 193 symptom free individuals has recently been established. The study has reported BME, synovitis and tenosynovitis (all with score≥1) to be present in 57.5%, 48.2% and 16.6% respectively [333]. Although only 28% had no single inflammatory-feature, scores of ≥2 were rare. Tenosynovitis was shown to be infrequent except for the extensor carpi ulnaris (ECU) tendon. Raw frequency data from this study is available and permits the use of MRI as potential biomarker.

4.2 Aims and overview

This programme of work aims to address the unmet research needs regarding MRI in an at-risk population. In a large cohort anti-CCP positive individuals with non-specific MSK symptoms this work aims to:

1) Describe the imaging characteristics demonstrated on MRI.
2) Determine whether MRI characteristics are associated with disease progression.

3) Evaluate the role of MRI in a predictive model for progression to inflammation arthritis.

4.3 Patients and Methods

4.3.1 Patients

A gadolinium enhanced MRI scan was offered to all eligible participants at their baseline CCP clinic appointment (at-risk cohort). Those with no contraindications to MR imaging attended the research unit for scanning of their dominant/most symptomatic hand and wrist. For this analysis, 98 CCP+ individuals were available. A large European database of MRI scans in symptom-free individuals (healthy controls) was available reporting the RAMRIS OMERACT scoring for joints, bones and tendons [333]. These authors have produced a reference range for abnormalities at specific joint per age category.

4.3.2 Clinical assessments

Clinical assessments were performed as previously described in Methods (Section 4.4 Protocol). All participants provided baseline demographic details, patient questionnaires, clinical history of symptoms and had a systems examination by a rheumatologist, (trained in the assessment of IA, [LH & CR]) including a joint count. Individuals then followed the study schedule as listed in Chapter 3.5 Protocol. Individuals attended 3 monthly visits for the first year and as clinically indicated thereafter or until they developed inflammatory arthritis; defined by the presence of at least one tender and swollen joint confirmed by a rheumatologist.

4.3.3 Imaging methods and scoring

4.3.3.1 Ultrasound assessments

Ultrasound assessments conducted as stated in Methods (Section 3.6.2.2). For this analysis, baseline ultrasound findings were limited to that reported in the MRI scanned hand. Presence of ultrasound power Doppler score ≥1 in any joint and greyscale score

≥2 in any joint were extracted from the database for this analysis. Flexor tendon scoring was available on a subset of the cohort (n=56).

4.3.3.2 MRI Assessments

The MRI protocol was followed as described in Methods (Section 3.6.2.3).

Images were anonymised and randomised prior to reporting. Scoring was performed independently by two musculoskeletal radiologists for synovitis, bone marrow oedema (BME) and erosions according to the OMERACT RAMRIS [303]. Digital flexor tendons were scored using a method in which the tendons were assessed from 1 cm proximal to 1 cm distal of each MCP joint [304]. Wrist tenosynovitis scores cover the tendons proximal to this using a previously described method [306]. Scoring of the anti CCP+ scans was undertaken in a case mix of healthy controls, RF+ arthralgia and early RA scans. Radiologists provided independently scored and consensus scoring. The consensus scores were used in this analysis. For consensus scoring, If any discrepancies occurred the radiologists reviewed images and a final score derived.

4.3.4 Statistical analysis

Frequency data reported by Magnus *et al* was available for the 193 MRI scans of healthy controls [333]. The authors reported that higher age was positively correlated with a higher total inflammation-score. The correlation was also demonstrated when considering each entity (erosions, tenosynovitis, BME, synovitis) separately. Hence, within this publication, an age category stratified reference range for each MRI imaging finding is available. Using this data, joint counts were created for each (erosion, synovitis, BME, tenosynovitis); a joint score was only positive and counted if <5% of age-matched healthy control in Magnus *et al*. had pathology present in that joint at that score level. This allowed age adjustment of scores calculated for MRI abnormalities.

Maximum MRI imaging scores observed per patient across all joints scored were initially trichotomised at 0, 1 and ≥2. Preparatory modelling indicated there was little difference

between scores of 0 and 1. As such, erosion, BME, synovitis and tenosynovitis scores were dichotomised at <2, ≥2 for regression analysis. This follows previous analysis evaluating ultrasound imaging of this cohort [237]. Continuous data was evaluated in the preliminary stages. However, dichotomisation facilitated the statistical modelling.

Following descriptive analysis of imaging characteristics of all individuals at risk to IA, the cohort was risk stratified according to the clinical model previously published [217]. This model included:

- Clinician-determined tenderness of small joints in the hands and feet (MCPs, PIPs, midtarsal and/or MTPs) (score 1 point),
- RF and/or anti CCP titre ≥3xULN (score 2 points),
- Early morning stiffness ≥30 minutes duration (score 1 point).
- Power Doppler ultrasound (score 1 point)

A total risk score of 0 indicated low risk, 1-2 moderate risk, 3-4 high risk.

Mantel-Haenszel tests of homogeneity (for the association between each of the MRI variables and progression to IA), were used to assess whether the association was the same irrespective of clinical risk of progression according to the clinical model.

Patient level analysis

Cox proportional hazards regression was used to ascertain the association between each MRI abnormality and time to IA. Modelling was performed using both total MRI abnormality scores and maximum score observed in any joint. The hazard ratios (HR) were adjusted for the clinical variables included in the risk score and additionally for ultrasound parameters; PD ≥1 and GS ≥2, which have previously been reported within this cohort [237]. Likelihood ratio tests were used to show whether each of the MRI variables improved model fit. Significance at $p<0.1$ was considered indicative of potential association with IA in this preliminary analysis. This significance level was selected in order to allow for all possible association to be considered given the sample size and exploratory nature of this work.

Joint level analysis

MCPs 2-5 and wrists scores were dichotomised as follows: ultrasound PD, ultrasound GS, MRI erosions, BME, synovitis and tenosynovitis (<2 or 2). At the wrist and MCP joints there are scores for multiple sites within each joint, the maximum score observed per joint was recorded. Cox proportional hazards models with standard errors adjusted for clustering of joints at the patient level were constructed, first on a univariable basis and then for each MRI feature adjusted for ultrasound GS and PD.

Binary logistic regression models were constructed to explore associations between MRI findings and clinical features at the joint level.

4.4 Results

4.4.1 Patient Characteristics

Ninety eight individuals from the CCP+ cohort were recruited to this study. Thirty percent (29/98) progressed to IA during the follow-up period. The baseline characteristics of the 98 individuals are presented in Table 4. Median time to progression was 31 weeks (IQR 24,67). Of those who did not progress, median follow-up was 132 weeks (IQR 75, 199).

Table 4 Baseline and ultrasound imaging characteristics of 98 individuals at risk of progression to IA

Characteristic		**CCP cohort (n=98)**
Age: mean (SD)		47.9 (12.2)
Female		69% (68)
SE	1 copy	45% (43)
	2 copies	22% (21)
ESR: median (IQR)		14.7 (11.6)
CRP	≥5 mg/dL	23% (22)

Anti-CCP	low positive	20% (20)
	high positive	80% (78)
RF	low positive	10% (10)
	high positive	36% (35)
Antibody (RF or CCP)	high positive	87% (85)
EMS	≥30 mins	32% (31)
Small joint tenderness	present	46% (45)
US power Doppler ≥2	present	9% (9)
US grey scale ≥2	present	23% (23)

All values are % (n) unless otherwise stated

Anti-CCP: anti-citrullinated cyclic peptide antibody, EMS: early morning stiffness, RF: rheumatoid factor, SD: standard deviation, SE: Shared epitope, US: ultrasound

In two individuals, both of whom did not progress to IA, no baseline ultrasound scan was available. The next available follow up scan results were used instead. Three patients did not receive gadolinium during their scan and therefore scores for tenosynovitis and synovitis were not available.

4.4.2 MR imaging findings in the at-risk cohort

Comparison to healthy controls

As described, MRI scores are adjusted using a healthy reference range. A joint/bone/tendon MRI finding was considered positive if the score obtained was present in <5% of age-matched healthy control population. For example, a synovitis score of 1 in MCP joint 2 in a 40 year old was considered negative if ≥5% of healthy controls in the same age category demonstrated this pathology. Tables 5-8 lists the values before and after adjustment. In a some cases this adjustment meant that scores were corrected to a lower score resulting in a greater number of individuals scoring 1, for example, than prior to adjustment (see Table 5 erosion adjusted scores for BMC1, trapezium and Table 8 tenosynovitis score for index).

Table 5 Number of individuals achieving MRI erosion score by location, before and after correction for age using healthy control data.

	Unadjusted erosion score (n=98)				Adjusted erosion score (n=98)			
Bone	0	1	2	≥3	0	1	2	≥3
MCPs								
MCP2	91	5	2	0	96	2	0	0
MCP3	92	5	1	0	98	0	0	0
MCP4	97	1	0	0	98	0	0	0
MCP5	93	3	2	0	96	2	0	0
WRIST								
BMC1	89	3	4	2	92	4	2	0
BMC2	92	5	0	1	92	5	0	1
BMC3	96	1	1	0	96	1	1	0
BMC4	97	1	0	0	97	1	0	0
BMC5	98	0	0	0	98	0	0	0
Trapezium	86	6	4	2	88	7	2	1
Trapezoid	89	7	1	1	96	1	0	1
Capitate	72	23	3	0	95	3	0	0
Hamate	90	7	1	0	92	6	0	0
Scaphoid	83	13	2	0	95	3	0	0
Lunate	81	12	4	1	93	4	1	0
Triquetrum	83	11	4	0	94	3	1	0
Pisiform	93	4	1	0	94	4	0	0
DRadius	91	5	2	0	91	5	2	0
DUlna	89	8	0	1	97	0	1	0

BMC base of metacarpal, MCP metacarpophalangeal, DRadius distal radius, DUlna distal ulna

Table 6 Number of individuals achieving MRI BME score by location, before and after correction for age using healthy control data.

	Unadjusted BME score (n=98)				Adjusted BME score (n=98)			
Bone	0	1	2	≥3	0	1	2	≥3
MCPs								
MCP2	96	2	0	0	97	1	0	0
MCP3	96	2	0	0	97	1	0	0
MCP4	96	2	0	0	96	2	0	0
MCP5	97	0	1	0	97	0	1	0
Wrist								
BMC1	90	6	0	2	94	2	1	1
BMC2	97	1	0	0	97	1	0	0
BMC3	98	0	0	0	98	0	0	0
BMC4	97	1	0	0	97	1	0	0
BMC5	98	0	0	0	98	0	0	0
Trapezium	88	7	2	1	88	7	2	1
Trapezoid	94	4	0	0	96	2	0	0
Capitate	87	10	1	0	87	10	1	0
Hamate	91	5	2	0	92	5	1	0

Scaphoid	90	8	0	0	98	0	0	0
Lunate	80	15	2	1	95	2	1	0
Triquetrum	92	3	3	0	95	0	3	0
Pisiform	97	1	0	0	97	1	0	0
DRadius	96	2	0	0	96	2	0	0
DUlna	92	6	0	0	98	0	0	0

BMC base of metacarpal, MCP metacarpophalangeal, DRadius distal radius, DUlna distal ulna

Table 7 Number of individuals achieving MRI synovitis score by location, before and after correction for age using healthy control data.

	Unadjusted synovitis score (n=95)				Adjusted synovitis score (n=95)			
Joint	0	1	2	≥3	0	1	2	≥3
MCPs								
MCP2	58	26	9	2	79	14	2	0
MCP3	70	24	1	0	93	2	0	0
MCP4	80	12	3	0	80	12	3	0
MCP5	69	18	7	1	74	17	4	0
Wrist								
DRUJ	62	31	1	1	90	4	1	0
ICJ	39	47	9	0	79	16	0	0
RCJ	35	46	13	1	77	15	3	0

MCP metacarpophalangeal, DRUJ distal radioulnar joint, ICJ intercarpal joint, RCJ radiocarpal joint

Table 8 Number of individuals achieving MRI tenosynovitis score by location, before and after correction for age using healthy control data.

	Unadjusted tenosynovitis score (n=95)				Adjusted tenosynovitis score (n=95)			
Tendon	0	1	2	≥3	0	1	2	≥3
Digital (flexor)								
Thumb*	90	4	1	0	-	-	-	-
Index	68	24	2	1	73	19	3	0
Middle	78	13	3	1	84	7	3	1
Ring	80	12	3	0	86	7	2	0
Little	81	13	1	0	81	13	1	0
Wrist (extensor and flexor)								
FCR	72	20	3	0	72	20	3	0
FPL	81	13	1	0	81	13	1	0
FDSFDP	71	22	2	0	71	22	2	0
FCU	92	3	0	0	92	3	0	0
I	85	7	3	0	85	7	3	0
II	86	8	0	1	86	8	0	1
III	89	3	3	0	89	3	3	0
IV	70	22	3	0	70	22	3	0
V	78	15	2	0	78	15	2	0
VI	55	28	10	2	76	17	1	1

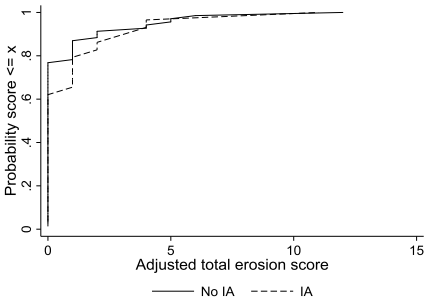
FCR flexor carpi radialis, FPL flexor pollicis longus, FDSFDP flexor digitorium superficialis and flexor digitorium profundus, FCU flexor carpi ulnaris, Extensor compartments I – extensor pollicis brevis and abductor pollicis longus, II extensor carpi radialis brevis and extensor carpi radialis longus, III extensor pollicis longus, IV extensor digitorium communis and extensor digitus proprius, V extensor digiti quinti porprius, VI extensor carpi ulnaris.
**Thumb was not included in paper by Mangnus et al and was not adjusted for age.*

Reviewing the adjusted versus unadjusted data several findings are evident. An unadjusted erosion score equalling 1 was relatively frequently reported in the carpal bones (particularly capitate, lunate and scaphoid). As this finding was also prevalent in the healthy controls the adjusted scores are lower. BME scores remained largely unchanged following adjustment, with the exception of the carpal bone scores which were adjusted from 1 to 0. Both synovitis and tenosynovitis were the most prevalent findings on MRI. Synovitis scoring ≥1 were reported in 48.2% of healthy controls in the Mangnus *et al* data. Subsequently, adjusted scores in this data set were reduced. In comparison, the adjusted scores for tenosynovitis demonstrated only minor variation as a consequence of the low prevalence of tenosynovitis scores in healthy controls.

Prior to adjustment, MRI scores of greater than 2, in an individual joint was reported in a total of 40 bones for erosions (40/1862, 2.1%), 15 bones for BME (15/1862, 0.8%), 48 joints for synovitis (48/665, 7.2%) and 42 tendons for tenosynovitis (42/1425, 2.9%). Following adjustment, an MR imaging score greater than 2 was reported in a total of 15 bones for erosions (15/1862, 0.8%), 12 bones for BME (12/1862, 0.6%), 13 joints for synovitis (12/665, 2%) and 31 tendons for tenosynovitis (31/1425, 2.2%). For all MRI findings an adjusted score of 0 for joints, bones and tendons was the most frequent.

Cumulative probability plots illustrate the total adjusted MRI scores for each MRI finding (Figure 7). The cohort has been divided into those who developed IA and those who did not (No IA). The individuals who developed IA had higher tenosynovitis and synovitis total scores at baseline imaging compared to the no IA.

a)



b)

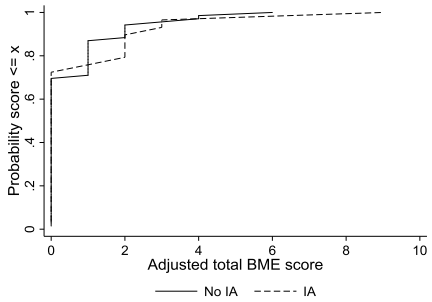
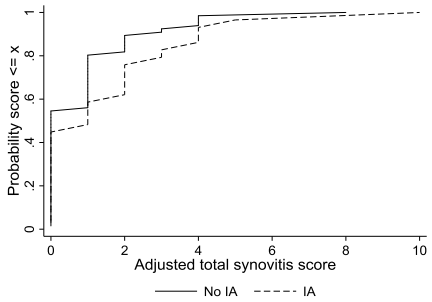


Figure 7 Cumulative probability plots of adjusted total MRI finding score for a) erosions, b) bone marrow oedema, c) synovitis, d) tenosynovitis.

c)



d)

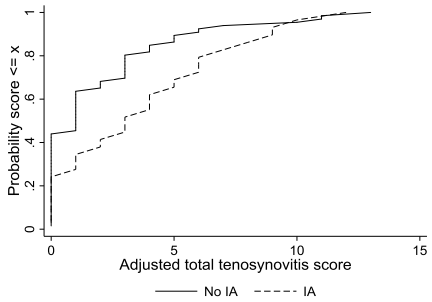


Figure 7 Cumulative probability plots of adjusted total MRI finding score for a) erosions, b) bone marrow oedema, c) synovitis, d) tenosynovitis.

MR imaging findings in at-risk individuals

Table 9 Maximum MRI imaging scores observed per patient across all joints.

MRI Imaging		**CCP Cohort (n=98)**
Erosion score	0	72% (71)
	1	18% (18)
	≥2	9% (9)
BME score	0	70% (69)
	1	19% (19)
	≥2	10% (10)
Synovitis score	0	52% (49/95)
	1	39% (37/95)
	≥2	9% (9/95)
Tenosynovitis score	0	38% (36/95)
	1	40% (38/95)
	≥2	22% (21/95)

Table 9 presents the maximum MRI scores per patient across all joints and summarises the main MRI baseline reporting. The adjusted scores by location are presented in Figure 8-11. (The data for score 0 is listed in tables 5-8 and is not presented in the figures.) Location of abnormalities were variable with some distinct patterns emerging. BME and erosions (scores ≥2) were reported in 10% and 9% of individuals respectively, with a preferential location of the carpal bones. In comparison, synovitis and tenosynovitis were seen more frequently, individuals scoring 1 reported in 39% and 40% respectively. Within at-risk individuals, synovitis score ≥1 were most commonly reported at MCP 5 (22.5%), radiocarpal joint (18.9%) and MCP 2 and intercarpal joint (16.8% for both). Synovitis was less commonly seen at MCP 3 and distal radioulnar joint (2% and 5% respectively). Only 9% (9/95) individuals had a synovitis score at any joint of ≥2. In comparison, tenosynovitis scores of ≥2 were reported in 22% of individuals (21/95). Tenosynovitis scoring ≥1 was reported in 25% of individuals at compartment IV (comprising of extensor digitorium communis and extensor digitus proprius), flexor digitorium superficialis/flexor digitorium profundus, and flexor carpi radialis tendons. Within the digital tendons, the index tendon had the highest reporting of abnormality (scores ≥1 in 23.1%). Figure 12 and Figure 13 show examples of tenosynovitis at the wrist and flexor tendons where the individual progressed to IA.

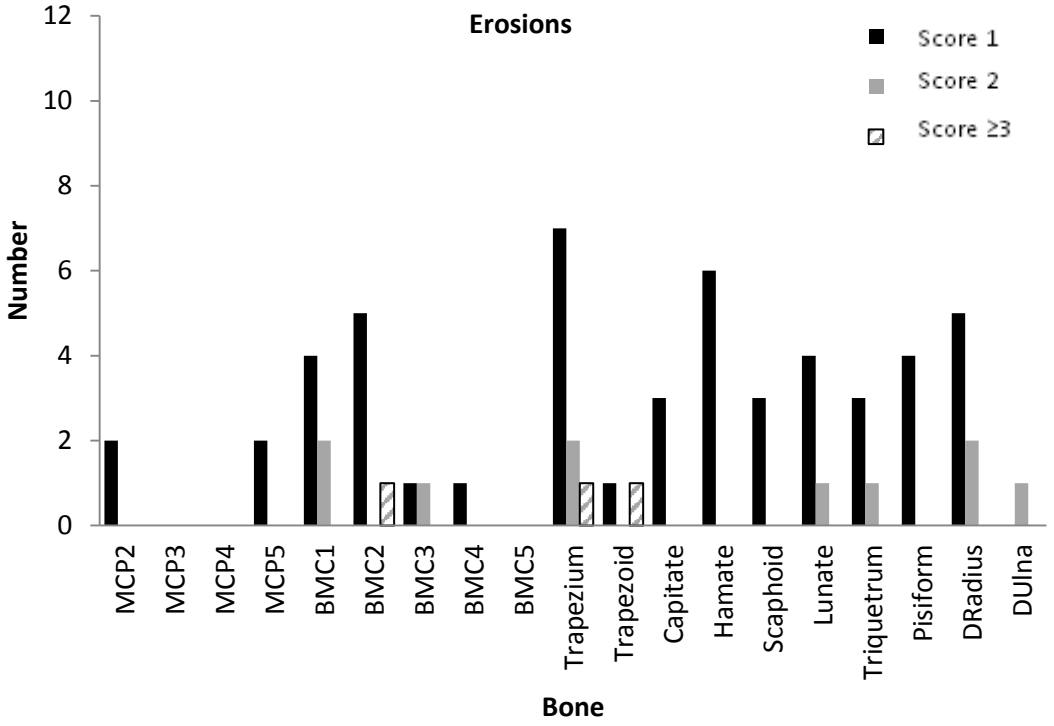


Figure 8 Number of individuals with adjusted MRI scores (1, 2 and ≥3) by location for erosions.

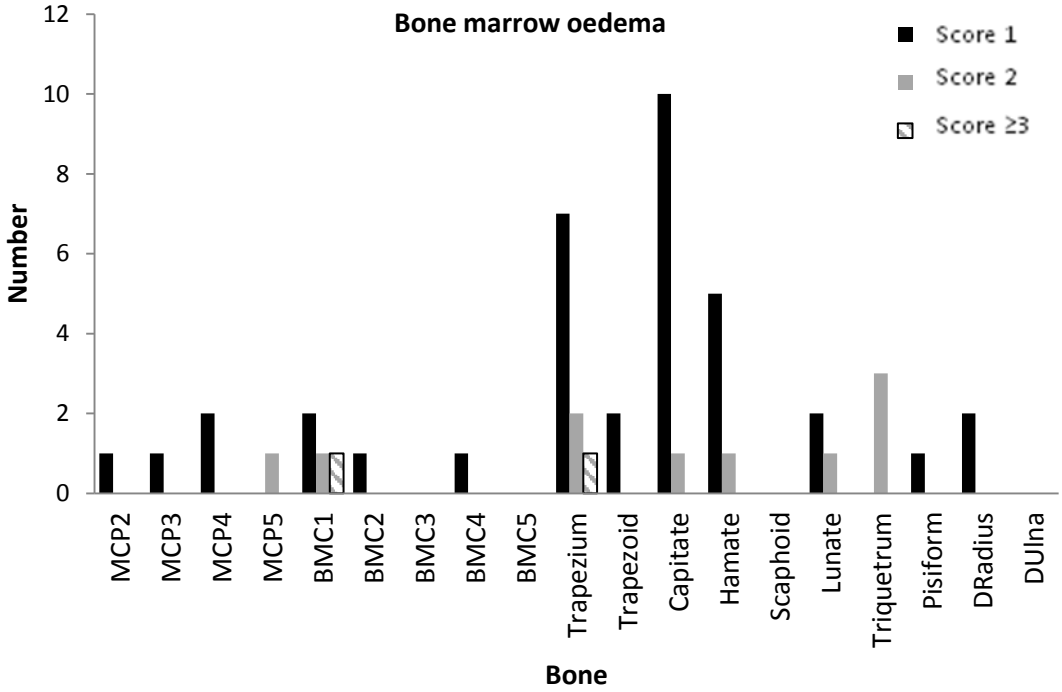


Figure 9 Number of individuals with adjusted MRI scores (1, 2 and ≥3) by location for BME.

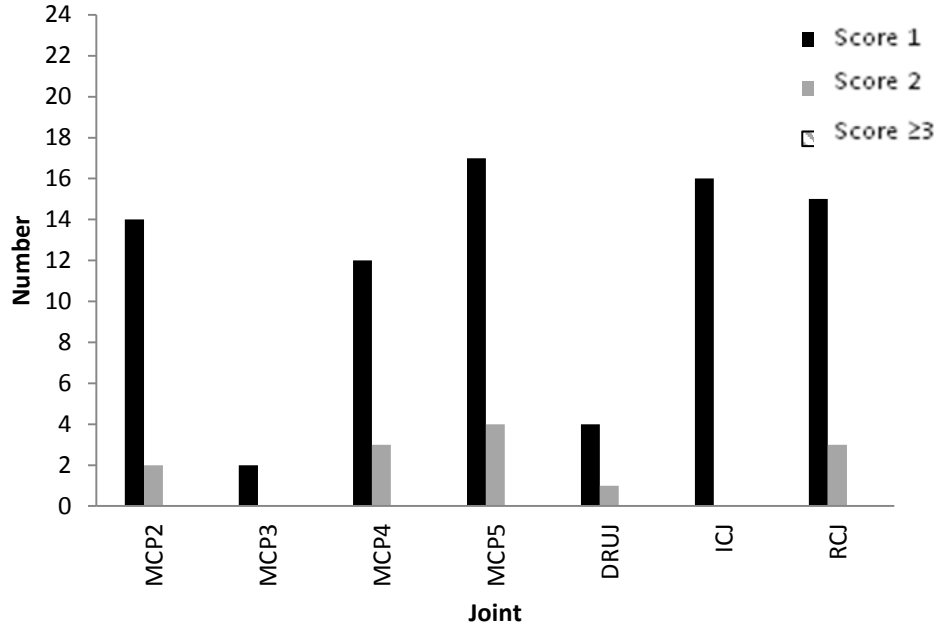


Figure 10 Number of individuals with adjusted MRI scores (1, 2 and ≥3) by location for synovitis.

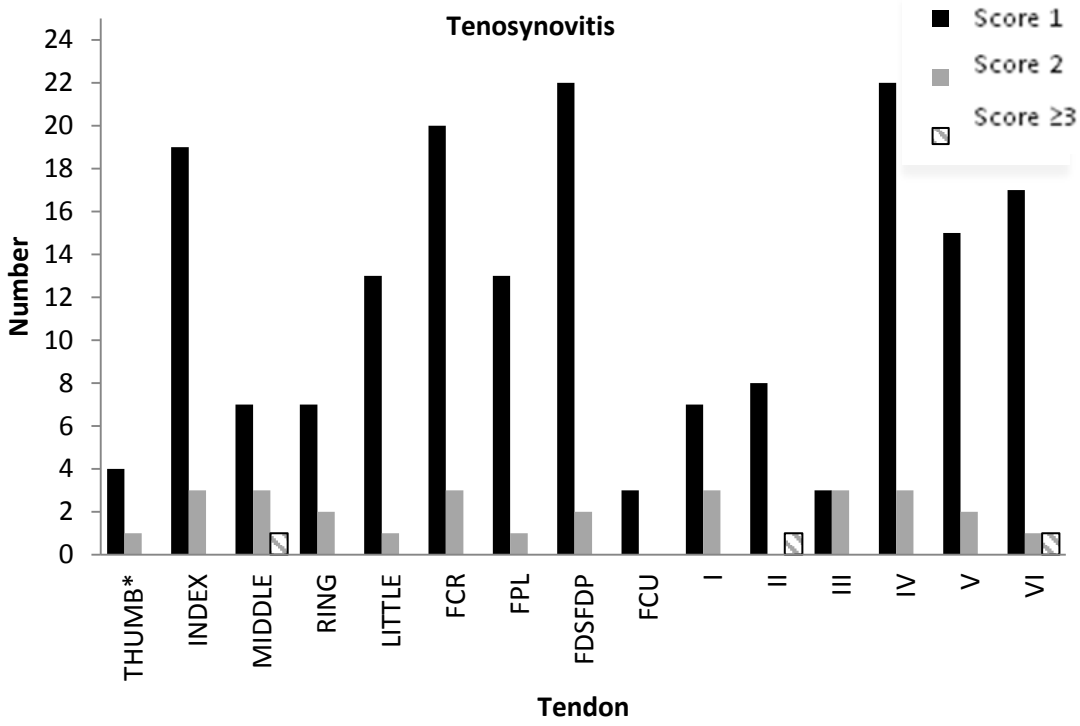


Figure 11 Number of individuals with adjusted MRI scores (1, 2 and ≥3) by location for tenosynovitis.

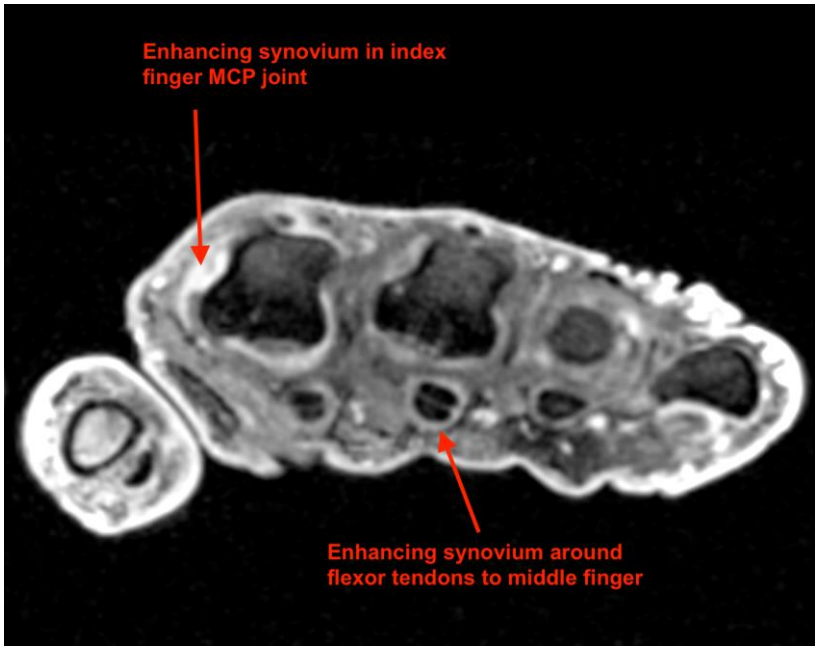
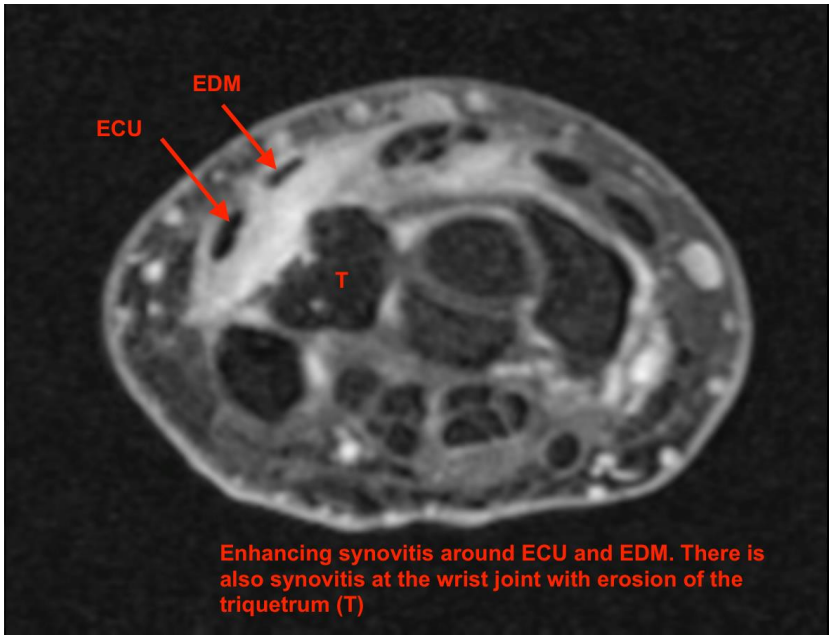


Figure 12 Axial reconstructions of Gadolinium enhanced 3D Spoiled Gradient Echo MRI image of the flexor tendons of an at-risk individual prior to the development of inflammatory arthritis.



(ECU extensor carpi ulnaris, EDM extensor digiti minimi , T Triquetrum)

Figure 13 Axial reconstructions of Gadolinium enhanced 3D Spoiled Gradient Echo MRI images of the wrist tendons of an at-risk individual prior to the development of inflammatory arthritis.

Risk stratification and MRI findings

The cohort was risk stratified using the previously reported clinical model (including ultrasound PD) [217] and Section 4.3.4 statistical analysis Table 10. Individuals with a MRI abnormality score ≥2 for each feature were compared in each level of progression according to whether or not they had progressed to IA (Figure 14). It was found that the moderate and high-risk categories demonstrate higher frequency of abnormalities ≥2 in the IA progression group compared to no IA. This was most evident when considering synovitis and tenosynovitis (Figure 14 c and d). The low risk category had no individuals with MRI scores ≥2 in either progression status, although the denominators are small. One individual progressed to IA in this strata and had no MRI abnormalities ≥2. Mantel-Haenszel test for homogeneity demonstrated that the effect of MRI is the same no matter which risk group an individual belongs (ranging from p=0.230 for BME to p=0.836 for tenosynovitis).

Table 10 Number of individuals with each MRI finding following risk stratification into low, moderate and high risk.

	Number with erosions		Number with BME	
Risk	No IA	IA	No IA	IA
Low	0/5	0/1	0/5	0/1
Mod	1/37	1/9	1/37	1/9
High	3/27	4/19	4/27	4/19

	Number with synovitis		Number with tenosynovitis	
Risk	No IA	IA	No IA	IA
Low	0/5	0/1	0/5	0/1
Mod	1/34	3/9	4/34	2/9
High	1/27	4/19	5/27	10/19

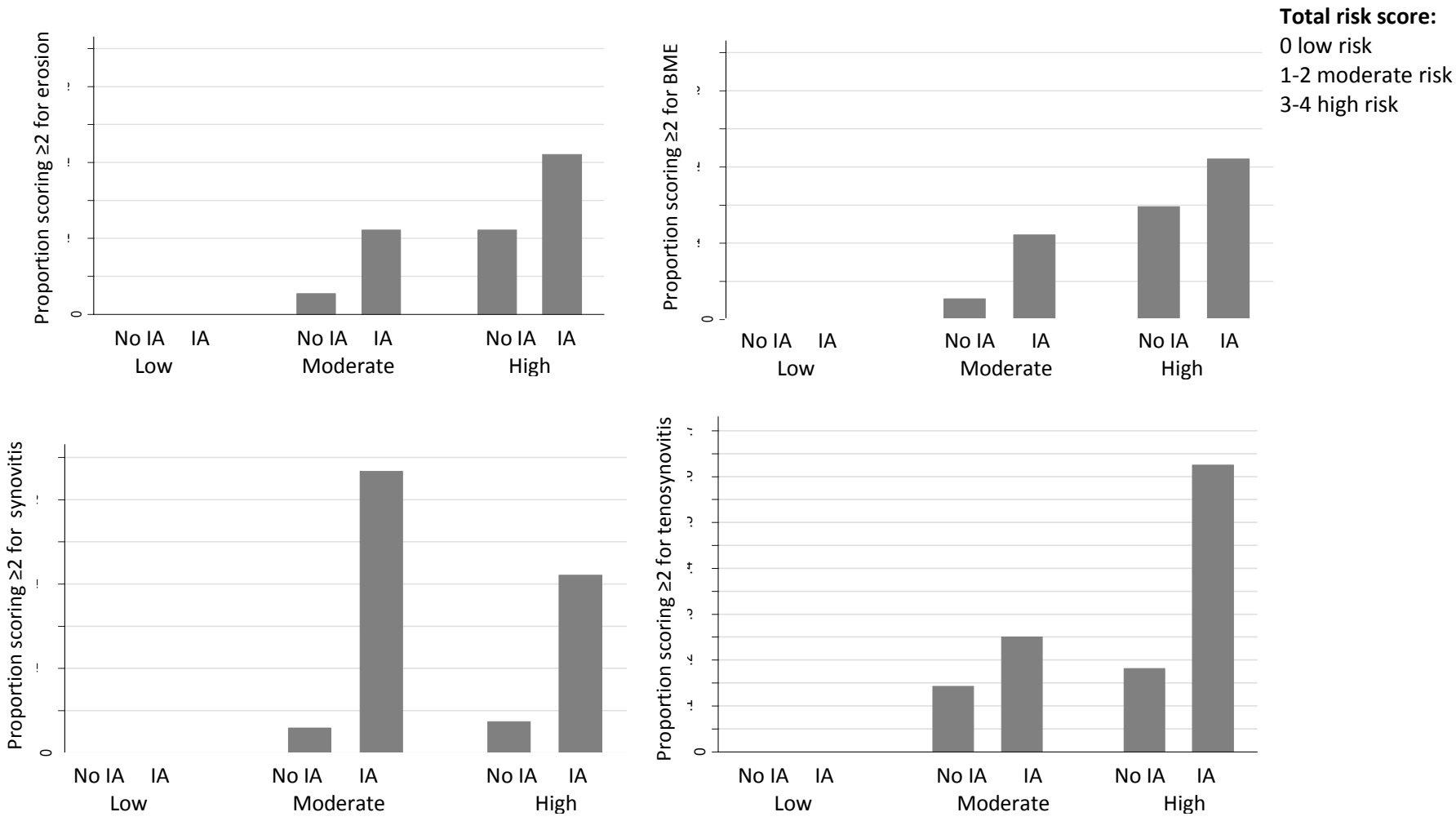


Figure 14 Risk stratification and MRI imaging: Proportions of individuals with score ≥2 according to progression status at each level of progression risk (a) erosions (b) BME (c) synovitis (d) tenosynovitis.

IA, Inflammatory arthritis

4.4.3 MRI (and ultrasound) findings and progression to IA

a) At the patient level:

Prior to adjustment, all ultrasound and MRI abnormalities were associated with risk of progression to IA. Ultrasound GS, ultrasound PD, MRI synovitis & tenosynovitis were statistically significant at $p<0.05$ and MRI erosion & BME at $p<0.1$ (Table 11).

Following adjustment controlling for the variables in original clinical model and for ultrasound features, the HR for MRI erosions, BME and synovitis decreased. HR for tenosynovitis was found to increase. In the adjusted analyses there was evidence that MRI-detected tenosynovitis was independently associated with time to IA (HR (90%CI)=3.89 (1.95, 7.76), p=0.001); Table 11. The hazard ratio for synovitis was substantive (HR (90%CI)=2.01 (0.92, 4.42), p=0.144) but, at this sample size the 90% confidence interval crossed 0. Ultrasound PD and GS remained substantive and statistically significant (at $p<0.1$) in adjusted models accounting for each of the MRI features. The clinical variables were not associated with progression in the model that included ultrasound variables and MRI tenosynovitis; small joint tenderness (1.34 (0.68, 2.65), p=0.478), high titre antibodies (1.07 (0.30, 3.79), p=0.930) nor EMS ≥30 minutes (1.37 (0.67, 2.82), p=0.473).

Table 11 Patient-level Cox proportional hazard modelling of associations between baseline MRI abnormalities and time to development of IA

					Adjusted* HR (90% CI) , *P* value			
Abnormality	**No IA (n=69) (%)**	**IA (n=29) (%)**	**Unadjusted HR (90% CI) , *P* value**	**GS & PD only**	**+erosion**	**+BME**	**+synovitis**	**+tenosynovitis**
US PD ≥2	1% (1)	28% (8)	7.57 (3.80, 15.09), p<0.001	3.93 (1.63, 9.47), p=0.010	3.93 (1.61, 9.57), p=0.011	3.66 (1.47, 9.09, p=0.019	3.14 (1.25, 7.88), p=0.040	5.38 (2.04, 14.17), p=0.004
US GS ≥2	12% (8)	52% (15)	4.71 (2.55, 8.71), p<0.001	2.51 (1.15, 5.49), p=0.053	2.51 (1.14, 5.52), p=0.056	2.55 (1.7, 5.58), p=0.049	2.45 (1.11, 5.41), p=0.064	2.38 (1.08, 5.26), p=0.073
MRI erosion ≥2	6% (4)	17% (5)	2.26 (1.01, 5.08), p=0.098		1.01 (0.39, 2.61),p=0.988			
MRI BME ≥2	7% (5)	17% (5)	2.40 (1.05, 5.47), p=0.080			1.31 (0.51, 3.36), p=0.632		
MRI synovitis ≥2	3% (2/66)	24% (7)	4.22 (2.03, 8.75), p=0.001				2.01 (0.92, 4.42), p=0.144	
MRI tenosynovitis ≥2	14% (9/66)	41% (12)	3.51 (1.89, 6.55), p=0.001					3.89 (1.95, 7.76), p=0.001

*Adjusted for presence of tenderness of small joints, EMS≥30 minutes, high antibody titre, maximum score MCPs, PIPs & wrists (bilateral) US GS ≥2 and US PD ≥2

When the analysis was repeated using total MRI abnormality scores instead of the maximum score observed in any joint, there was no change in the conclusion that of the MRI variables, only tenosynovitis was independently associated with risk of progression. Similarly, ultrasound GS and PD were always independently associated with risk of progression to IA.

b) At the joint level:

As MRI was performed on the most symptomatic/dominant hand (with scoring at wrists and MCPs 2-5), the presented data represents the joints that progressed that were scored. A total of 490 joints are considered. Of these, 29 progressed to clinical synovitis.

Prior to adjustment MRI BME, synovitis and tenosynovitis were all associated with risk of that specific joint progressing to clinical synovitis (CS) (HR 5.63, 4.45 and 5.96 respectively, all p=<0.05) Table 12. Of the ultrasound parameters, both GS and PD were also associated with joint progression. MRI erosions were the only imaging pathology not to be associated with risk of joint progression (HR 2.82 90% CI 0.38, 20.94, p=0.310).

When adjusted for the ultrasound variables previously demonstrated to be associated with progression (ultrasound PD≥2 and for GS≥2), HRs for MRI BME increased, Table 12. There was evidence that BME and tenosynovitis were associated with the risk of progression to clinical synovitis (HR=6.16, p=0.011 and HR=7.22, p<0.001 respectively). The adjusted HR for MRI erosions was substantive but not significant, HR 2.5, p=0.171. In contrast, the HR for MRI synovitis decreased.

Table 12 Joint-level Cox proportional hazard modelling of associations between baseline MRI abnormalities and time to development of clinical synovitis (MCPs 2-5 and wrists only)

				Adjusted* HR (90% CI) , *P* value				
Abnormality	**No CS (n=466) (%)**	**CS (n=24) (%)**	**Unadjusted HR (90% CI) , *P* value**	**GS & PD only**	**+erosion**	**+BME**	**+synovitis**	**+tenosynovitis**
US PD ≥2	<1% (3)	4% (1)	7.09 (1.08, 46.54), p=0.087	1.51 (0.14, 16.32), p=0.776	1.33 (0.20, 8.72), p=0.801	1.44 (0.13, 16.16), p=0.806	1.71 (0.15, 19.58), p=0.717	1.23 (0.24, 6.34), p=0.832
US GS ≥2	3% (14)	21% (5)	8.40 (4.08, 17.29), p<0.001	6.98 (3.09, 15.80), p<0.001	7.23 (3.28, 15.93), p<0.001	7.81 (3.39, 18.00), p<0.001	5.77 (2.40, 13.87), p=0.001	6.79 (2.91, 15.85), p<0.001
MRI erosion ≥2	2% (7)	4% (1)	2.82 (0.53, 15.18), p=0.310		2.50 (0.83, 7.51), p=0.171			
MRI BME ≥2	2% (8)	8% (2)	5.63 (1.82, 17.37), p=0.012			6.16 (1.91, 19.90), p=0.011		
MRI synovitis ≥2	2% (10/451)	8% (2)	4.45 (1.51, 13.11), p=0.023				2.20 (0.67, 7.17), p=0.274	
MRI tenosynovitis ≥2	4% (17/451)	25% (6)	8.46 (4.21, 16.98), p<0.001					7.22 (3.66, 14.25), p<0.001

*CS = Clinical synovitis, *Adjusted for presence of tenderness of small joints, EMS≥30 minutes, high antibody titre, US PD≥2 and US GS≥2*

In the adjusted models, ultrasound GS remained significant while the adjusted HRs for ultrasound PD were close to 1.

4.4.4 Association to clinical variables

Table 13 and Table 14 demonstrate the logistic regression analysis for physician and patient reported joint tenderness with respect to imaging findings (MRI and ultrasound). All variables were dichotomised to >1 since none of the tender joints had score >2 for ultrasound PD, MRI erosions and BME. The results indicate that no imaging modality or finding was significantly associated with physician determined joint tenderness. Patient reported joint tenderness was recorded as any recent pain at fingers/hand and wrist. The regression analysis indicates that ultrasound PD was associated with patient reported tenderness (HR 3.95, 90% CI (1.33, 11.69), p=0.037). It is noted that PD signal was only detected in a minority of painful joints. The remaining imaging variables were not associated. This analysis indicates there is limited evidence to associate MRI findings with joint tenderness, either physician or patient reported.

Table 13 Results of logistic regression models of physician-determined tenderness at the joint level.

Abnormality		**Not tender n=446**	**Tender n=44**	**Unadjusted HR (90% CI) , *P* value**	**Adjusted HR (90% CI) , *P* value**
US PD	≥1	3% (14)	5% (2)	1.47 (0.41, 5.30), p=0.621	2.09 (0.67, 6.57), p=0.289
US GS	≥1	18% (82)	16% (7)	0.84 (0.34, 2.07), p=0.750	0.77 (0.33, 1.80), p=0.613

MRI erosion	≥1	6% (28)	7% (3)	1.09 (0.44, 2.73), p=0.874	1.44 (0.48, 4.33), p=0.579
MRI BME	≥1	6% (28)	7% (3)	1.09 (0.42, 2.81), p=0.878	1.07 (0.33, 3.49), p=0.928
MRI synovitis	≥1	18% (76/431)	9% (4)	0.47 (0.19, 1.14), p=0.159	0.46 (0.20, 1.07), p=0.131
MRI tenosynovitis	≥1	23% (97/431)	16% (7)	0.65 (0.30, 1.41), p=0.360	0.68 (.29, 1.58), p=0.452

Table 14 Results of logistic regression models of patient-reported pain at the joint area level.

Abnormality		**Not painful n=115**	**Painful n=71**	**Unadjusted HR (90% CI) , *P* value**	**Adjusted HR (90% CI) , *P* value**
US PD	≥1	3% (4)	13% (9)	4.03 (1.49, 10.89), p=0.021	3.95 (1.33, 11.69), p=0.037
US GS	≥1	32% (37)	37% (26)	1.22 (0.74, 2.00), p=0.512	0.95 (0.54, 1.68), p=0.892
MRI erosion	≥1	17% (19)	15% (11)	0.93 (0.46, 1.87), p=0.858	0.84 (0.36, 1.99), p=0.741
MRI BME	≥1	17% (20)	15% (11)	0.87 (0.44, 1.73), p=0.740	0.84 (0.38, 1.85), p=0.714
MRI synovitis	≥1	30% (33/109)	31% (22)	1.03 (0.59, 1.81), p=0.921	0.91 (0.51, 1.60), p=0.775
MRI tenosynovitis	≥1	40% (44/109)	51% (36)	1.52 (0.90, 2.56), p=0.186	1.57 (0.91, 2.70), p=0.174

4.4.5 Performance of ultrasound and MRI for tenosynovitis scoring

MRI tenosynovitis scores included wrist and digital tendons, whereas ultrasound scoring was restricted to the flexor tendons. For comparison between modalities the analysis is therefore limited to findings reported at the digital tendons. Fifty seven percent (56/98) of individuals in the CCP+ cohort had complete ultrasound tendon scores. A total of 224 tendon scores were available for comparison. None of these tendons had an ultrasound PD score >0 and therefore ultrasound GS only is compared.

The cross-tab table below illustrates the MRI score and ultrasound GS score for the 224 tendons. Eighty eight percent agreement was seen between the two modalities (prevalence-adjusted, bias-adjusted Kappa 0.77). However, this was largely due to agreement in the reporting of the absence of tenosynovitis (score 0). Only 2.6% (6/224) of tendons had a ultrasound GS score >0, whilst MRI detected tenosynovitis >0 was found in 10.3% (23/224). Considering each category individually, agreement was 94% for scores of 0 and just 19% for scores >0. This suggests that MRI detects more tenosynovitis, or alternatively is more sensitive.

Table 15 Comparing MRI tenosynovitis score to ultrasound GS score

	Ultrasound GS score			
MRI score	0	1	2	Total
0	195	3	0	198
1	20	2	0	22
2	3	0	1	4
Total	218	5	1	224

In an unadjusted analysis of this restricted patient group; MRI tenosynovitis (of flexor tendons only) was still predictive of future development of clinical synovitis (HR (90% CI)=18.3 (3.5, 95.8), p=0.004). None of the 8 joints that developed clinical synovitis had ultrasound GS scores of >0, whilst 6 of the 220 of the joints that did not develop clinical synovitis did have ultrasound GS scores >0.

Table 16 Restricted cohort MRI scoring and development of clinical synovitis (CS)

	Restricted patients and joints	
	No CS	CS
MRI score	n=216	n=8
0	89% (193)	63% (5)
1	9% (20)	25% (2)
≥2	1% (3)	13% (1)

4.5 Discussion

In RA, MR imaging of patients with clinical synovitis demonstrates significant pathology, which is predictive of future disease progression and joint damage [317, 334-337]. There are limited data available regarding MR findings of at-risk individuals without detectable clinical synovitis. In this chapter, the MRI findings of 98 CCP+ individuals from the Leeds at-risk clinic are reported. This is the largest analysis of MRI findings in an at-risk population.

Contrast enhanced MR images obtained from a 3.0 Tesla scanner have facilitated the RAMRIS scoring of all four MRI variables: erosions, BME, synovitis and tenosynovitis. This is in contrast to previous studies which have failed to comment on erosions [325] and tenosynovitis [233, 234]. The moderate levels of inflammatory pathology reported in healthy control populations necessitates the adjustment of RAMRIS scores to abnormalities seen in healthy controls [332]. Following adjustment, whilst a proportion of the synovitis and erosion scores were within the normal range, less correction was required in tenosynovitis and BME scores. Overall, maximum scores ≥2 for each MRI finding were reported in a small proportion of individuals.

MRI synovitis was most prevalent in the wrist joints similar to that reported in the CSA population. In contrast, MCP 5 joint had greater reports of synovitis compared to the other MCP joints. The carpal bones had the highest reports of BME and erosions. Tenosynovitis had preferential location for flexor tendons of the wrist and extensor tendon compartments VI, V and IV which is similar to that reported in CSA [324]. The

index finger had the highest reports of tenosynovitis and MCP3 (middle finger) one of the least, which is in contrast to findings in the CSA population. Cross study comparison is possible due to both correcting to the same healthy control data. However, it is important to note that the study populations differ with regards to autoantibody status and therefore potential risk; 16% of the CSA being ACPA positive as opposed to 100% in the Leeds cohort.

Previous studies have attempted to determine the significance of MRI findings in similar at-risk populations. Unfortunately, sample size had limited reporting on associations to progression [233, 234]. In this cohort, adjusting for other imaging variables, MRI tenosynovitis has been shown to predict future development of IA in at-risk individuals (HR 3.89, p=0.001). Furthermore, the joint level findings indicate that MRI BME and tenosynovitis predicts development of clinical synovitis in individual joints (HR=6.16, p=0.011 and HR=7.22, p<0.001 respectively). This reflects findings from a recently published study in a much smaller cohort of 20 ACPA+ individuals [235]. Here, tenosynovitis was reported in 80% (16/20) of the cohort and is importantly absent in their healthy control population. On an exploratory level, the authors were able to report that presence of tenosynovitis in 2 or more sites was associated with later development of IA. Results reported in this chapter verify findings from Kleyer *et al* and provide further evidence that tenosynovitis is one of the earliest features reported on MR imaging in those at risk to IA progression.

Further stages of multivariable modelling were limited due to study size and lack of individuals that have progressed to IA. However, the clinical utility of MRI findings are illustrated through both the Cox proportional hazard regression analyses and the risk stratification exercise. In the former, the clinical variables were not associated with progression in the model that included ultrasound variables and MRI tenosynovitis. At the patient level, MRI tenosynovitis association with risk to IA progression is comparable to ultrasound findings, whilst at a joint level there is additional benefit. It is suggested that MRI is a feasible or even a preferred option in some units where experienced MSK ultrasonographers may be lacking. In the risk stratification exercise, the original model [217] was applied. Within each strata of risk, the proportion of individuals with MRI

score ≥2 as per progression status was presented. It was hypothesised that MR imaging findings may be particularly useful in those individuals at low risk for progression. However, the denominators for these groups were very small; there were only 6 individuals at low risk and just 1 progressed to IA. This one patient did not have any MRI abnormalities scoring ≥2. The test of homogeneity suggested that MRI findings were strongly associated with progression in all strata. As illustrated in Figure 13, MR imaging could be used as an adjunct to previous models to stratify individuals. With continual recruitment and follow-up of existing participants, it is anticipated that a larger risk stratification exercise involving all parameters would assist in the development of combined model.

The ability of both ultrasound PD and GS has been reported to be predictive of IA development [237]. However, within this analysis, although both imaging variables remained significant at a patient level, only GS seemed to be associated with the risk of progression to CS at the joint level. This was an unexpected finding. It is noted that in this study, ultrasound analysis is limited to unilateral hand as opposed to the previously bilateral hands finding. Furthermore, very few of the joints were found to have PD≥2 in unilateral MCPs2-5 and wrist. This may account for some of the findings. However, repeating the analysis using PD≥1 did not affect the outcome. It is recommended that a larger cohort should be analysed for clarification.

The association to progression with MRI tenosynovitis calls into question whether these features can be detected on ultrasound. Not all tendon compartments of the wrist are routinely imaged in ultrasound and the protocol for this study was only recently changed to include flexor tendon in the baseline ultrasound assessment. The direct comparison between the two modalities is therefore limited. However, results suggest that MRI reports greater frequency of tenosynovitis compared to ultrasound. Furthermore, in this limited analysis, MRI tenosynovitis was associated to development of clinical synovitis in a joint, in contrast to ultrasound GS tenosynovitis. This finding correlates with other studies which have compared modalities in the identification of tenosynovitis [338].

As discussed in the introduction (Section 2.1.2.1), there has been increased interest surrounding the pathogenic role of ACPA and osteoclast activation with subsequent bone damage [49]. The presence of BME and erosions detected by MRI in at-risk individuals offers an opportunity to explore potential relationships. Bone mineral density has been assessed with respect to MRI detected inflammation in the CSA cohort [339]. The authors demonstrated that severe bone loss was associated with MRI inflammation and progression to IA. It is difficult to draw to many conclusion from this as only 13% (14/108) of individuals were ACPA positive. Reviewing the results from the adjusted data of this CCP+ cohort, it is apparent that just under a third of individuals had evidence of either BME or erosions. Thus, the presence of ACPA does not equate to MRI pathology. BME was associated with progression to clinical synovitis within a joint but not an individuals' risk of progression. This does suggest that BME is an important finding in the joints of at-risk ACPA+ individuals. It was hypothesised that BME and joint damage would correlate with joint tenderness. In this study there is no evidence to suggest BME or indeed any other MRI finding is associated with patient or physicians reported joint tenderness. Patient reported joint tenderness was collected through questionnaires which recorded location of pain experienced in the last 7 days. This may not be a true reflection of pain pattern in the preceding months. However, very few of the tender joints had any detectable pathology. This limits the plausibility of strong association between MRI features and tenderness in the majority of these patients. Ultrasound PD did have a substantive HR 2.09 (90% CI 0.67, 6.57) which therefore supports further analysis of joint tenderness and imaging findings in a larger sample size.

Cost, duration of examination, comfort and accessibility are still the main barriers to the widespread application of MRI in clinical practice [340]. In the research setting, these images are RAMRIS scored by radiologists trained in musculoskeletal radiology. Whether this is feasible in all clinical departments is questionable and may limit the transferability of this modality into common practice. However, this study has demonstrated a clear benefit of MR imaging in evaluating individuals at risk of IA development.

4.6 Limitations

This study has established that MRI findings are associated to risk of progression in at-risk individuals. However, further multivariable modelling to consider all variables would require a larger data set. It is hoped that this will be achievable with ongoing recruitment.

Healthy control data has been provided by a large European database [333]. It is acknowledged that the data originates from different MRI scanners and RAMRIS scoring from other scorers. It was felt that despite these discrepancies, using healthy control data to correct scores allowed for a robust analysis.

The presented analysis was restricted to imaging of the most symptomatic or dominant hand. It was considered imperative to first focus on features reported in the hands following the ultrasound analysis of 136 individuals from the same cohort. Here, ultrasound GS reported at the metatarsal phalangeal (MTP) joints was less discriminating at differentiating between progressors and non progressors [237]. MRI is able to provide detailed evaluation of MTP which may follow a similar pattern of abnormality as described in the ultrasound analysis. van Steenbergen *et al* reported on the MRI features at the feet in the CSA cohort [324]. Whilst, low grade synovitis and BME (scores of ≥1) was reported in a fifth of individuals, scores ≥2 were infrequent. In the combined inflammation score (sum of tenosynovitis, synovitis and BME), MTP joint location had the lowest scores. Furthermore, scanning extra regions would require longer scan time and there are time implications to both the individual scanned and the radiologist reporting images. Restricting the imaging assessment to the hand is thought to be a more feasible option for assessment of relevant inflammation.

As alluded to earlier, it is challenging to compare these findings with other studies in the field since only one other study has corrected for changes found in health controls [324,

325]. However, in this population of at-risk, the CSA cohort, under a third were autoantibody positive and hence drawing comparisons is limited. Specifications of MRI scanners, radiologists/specialists scoring and changes to the OMERACT RAMRIS scoring for tenosynovitis are all important factors to consider when comparing across studies [304].

4.7 Key points

- MRI tenosynovitis at initial assessment independently predicts likelihood to IA in individuals with systemic autoimmunity and MSK symptoms.
- At a joint level, MRI tenosynovitis and BME are associated with the risk of progression to clinical synovitis.
- There is limited evidence to suggest that MRI findings are associated with joint tenderness, either physician or patient reported.
- MRI reported tenosynovitis appears to be more sensitive at detecting tenosynovitis compared to ultrasound, although further validation is required.

4.8 Conclusions

This study has presented the largest MRI data set for an autoantibody positive at-risk cohort. The correction of the data to a healthy control reference range permitted the clinical utility of MR imaging as a viable biomarker to assist in evaluation of risk. Baseline MRI characteristics, following adjustment, indicated that tenosynovitis was the most reported finding in these at-risk individuals. The levels of MRI detected inflammation and bone change was otherwise relatively modest. However, MRI findings, particularly tenosynovitis, provided predictive capability over and above the variables included in the clinical model and ultrasound GS & PD. There has been no analysis considering ultrasound determined tenosynovitis in this cohort due to a limitation in the dataset numbers. Further clarification is required as to the benefit of MRI detected tenosynovitis compared to ultrasound detected tenosynovitis. This may have

implication on the future management pathways of at-risk individuals and will therefore be an important focus of future work.

In this cohort of anti CCP+ individuals, there is insufficient evidence to suggest an association between tender joints and MRI features. However it is acknowledged that the current study had limitations in the data collection of painful joints.

Conclusions drawn from this study confirm preliminary pilot data. MRI in at-risk individuals has been confirmed as an effective imaging biomarker for predicting development of arthritis. These results supports the inclusion of MRI markers, in particular tenosynovitis, in future risk stratification models. It is recommended that larger sample size studies be conducted to confirm the conclusions from this study.

5. MicroRNA profiling of matched samples in individuals with systemic autoimmunity and arthralgia who progress to RA

In this chapter, the potential of microRNAs (MiRNAs) as a biomarker in identifying individuals at risk of RA is explored. Basic science techniques were applied to a clearly defined cohort to identify possible pathogenic mechanisms and specifically the clinical utility of miRNAs.

5.1 Introduction

MiRNAs are a highly conserved class of short non-coding RNAs (21-25 nucleotides) that serve as transcriptional negative regulators, involved in fine tuning of genes involved in cell differentiation, metabolism and immunity [82, 83]. Within the nucleus, miRNAs are transcribed by RNA polymerase II into molecules called primary miRNAs (pri-MRNAs) (Figure 15). They are then cleaved in association with a RNA binding protein (labelled DGCR8) by a RNase III enzyme (Drosha) to form precursor miRNAs (pre-miRNAs). Pre-miRNAs are characterised by a stable hairpin like structure. Once exported to the cytoplasm, the pre-miRNA undergo further cleavage by another RNase III enzyme to form unstable double stranded miRNA duplexes. The unwound miRNA can then enter the RNA-induced silencing complex (RISC) where it induces a gene silencing effect on transcribed messenger RNA (mRNA).

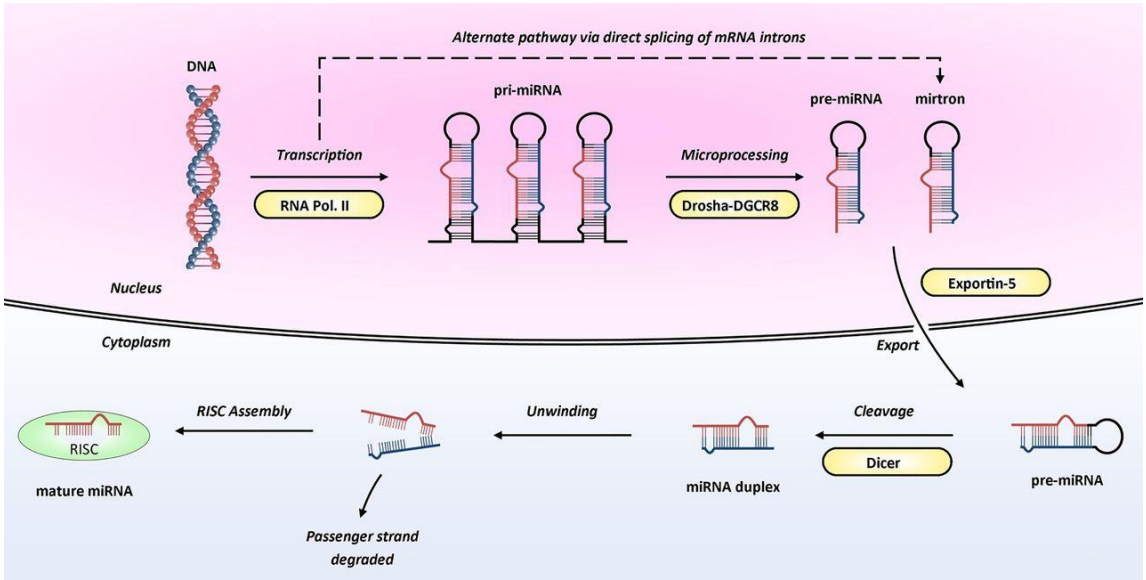


Figure 15 Schematic of the biogenesis of miRNA

Note the dashed line represents an alternative pathway in which pre-miRNA can be formed form direct splicing of introns these are termed as mitrons. Reprinted by permission from BMJ Publishing Group Ltd & British Cardiovascular Society [341]. Copyright 2015

The mode of action of miRNAs are presented in Figure 16. There are 3 main endpoints for mRNA. In the absence of miRNA, protein translation occurs (A). In the presence of miRNA with a 'near-perfect' complementarity, the miRNA binds (usually in the 3' untranslated regions of mRNAs) and represses translation, effectively 'silencing' the gene, inhibiting protein synthesis (B). When there is a perfect complementarity, the miRNA binding inhibits protein synthesis through the induction of mRNA degradation.

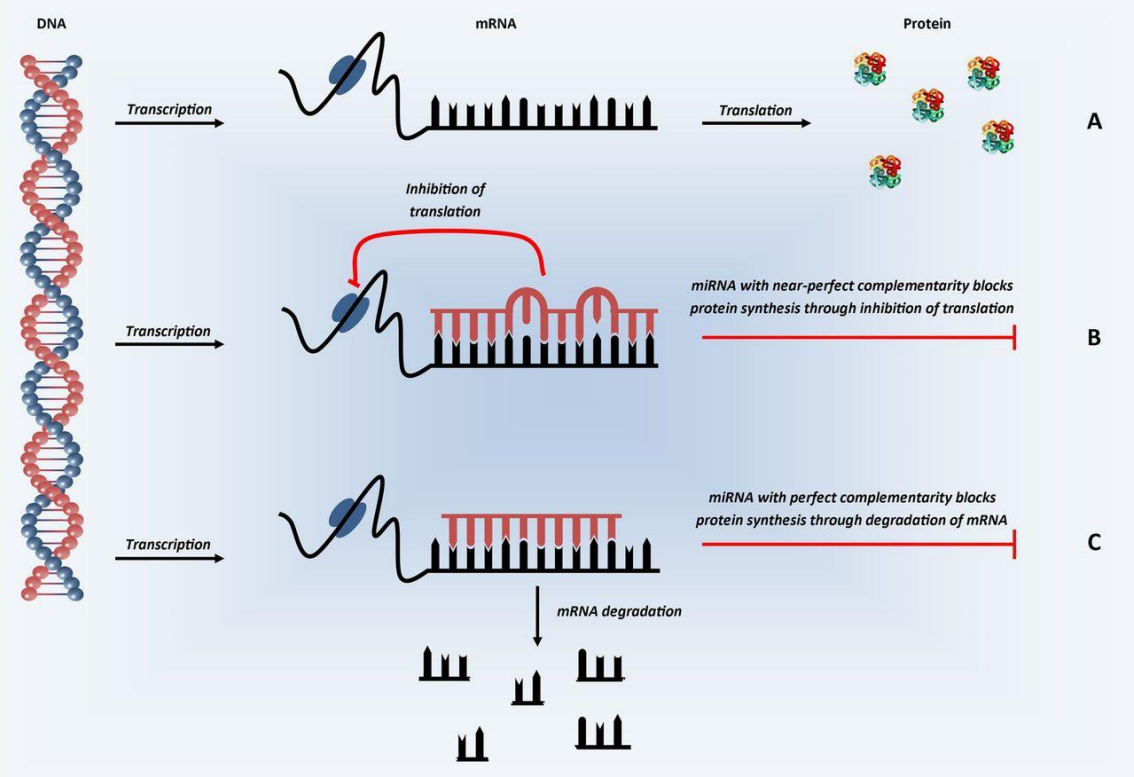


Figure 16 Mode of action of miRNA (simplified schematic)

Reprinted by permission from BMJ Publishing Group Ltd & British Cardiovascular Society [341]. Copyright 2015

Around the time of their discovery in the 1990s, miRNAs were thought to have little importance in cellular processes. However, within 10 years, advances in epigenetics enabled miRNAs to be recognised as pivotal regulators of gene expression, regulating at least 30% of mRNA [84, 85]. They have significant influence over diverse biological activities including apoptosis, immune functions, development, metabolism and

proliferation. Over 2000 miRNAs have now been identified [342]. Specific targets are known in only a small proportion, although their emerging role in autoimmune disease is becoming apparent [343-346].

Circulating miRNAs have been shown to have diagnostic and prognostic potential in malignancies [347-349]. Interest in miRNAs as a candidate biomarker for other diseases is based upon; (1) relative stability in the circulation and (2) detection in a wide range of readily accessible biologic sources (including serum and plasma) [350-353]. Consistency and reproducibility of results are even seen when frozen samples were investigated [354].

Research in the field of RA, has reported several miRNAs to be dysregulated, possibly suggesting a pathogenic role [355]. Initially studies focused on expression from synovial fibroblast cells. Dysregulation was reported in two key miRNAs; miR-146a, miR-155 [88, 89, 356]. Subsequently, further miRNAs have been reported. The well characterised miRNAs, patterns of dysregulation and proposed function are presented in Table 17.

Table 17 MiRNAs found to be dysregulated in patients with RA

miRNA	Pattern of dysregulation	Proposed function
miR-16	▸ Upregulated in RA PBMCs, SF level lower than plasma levels in patients with RA.	Regulate TNF-α signalling. [356, 357]
miR-18a	▸ Upregulated in RASFs.	Increase in MMPs and mediators of inflammation via NF-κB signalling. [358]
miR-21	▸ Upregulated in RA plasma, downregulated in PBMCS, PB $CD4^+$ cells & SF $CD4^+$	Maintaining balance between immune activation and tolerance regulation - Th17/Treg imbalance [359]
miR-22	▸ Downregulated in RA ST.	Inhibits CYR61. CYR61 stimulates RASF proliferation & IL-6 secretion – thus promotes synovial tissue hyperplasia) [360]
miR-34a	▸Downregulated in RASFs.	Regulator of cell death. Contributes to impaired apoptosis of RASFs. [361]
miR-132	▸ Upregulated in RA PBMCs, SF level lower than plasma levels in patients with RA.	Regulation of T lymphocytes & TNF-α production. [356, 357]

miR-146a	▸ Upregulated in RA ST, RASFs and RA PBMCs, Upregulated in RA $CD4^+$ from PB and SF. Downregulated in serum/plasma	Regulation of TNF-α. Th1 and Treg homeostasis [88, 89, 356, 357, 362]
miR-155	▸ Upregulated in RA ST/SF, RASFs, PBMCs and whole blood. Downregulated in serum/plasma	Regulation of cytokines including TNF-α, IL-6,IL-8, IL-1β and IL-10. Homeostasis of Treg cells and IL-17 producing cells. [89, 363]
miR-203	▸Upregulated in RASFs.	Increased secretion of MMP-1 and IL-6 via the NF-κB pathway. [87]
miR-223	▸ Upregulated in RA $CD4^+$ & RASFs, Abundance lower in SF than in plasma of patients with RA.	Involvement in osteoclastogenesis. [357, 364, 365]

(PB peripheral Blood, RASF Rheumatoid arthritis synovial fibroblasts, SF synovial fluid, ST synovial tissue.)

Original data from Wittmann J et al, permission from BMJ Publishing Group Ltd [355] Copyright 2011. Table updated with current literature regarding proposed function and additional miRNAs.

Other miRNAs that have been reported in RA include miR-23b, miR-30a, miR-24, miR-26a, miR-124a, miR-125a-3p/-5p, miR-125b/ miR-99a/ miR-100, miR-126-3p, miR-133a, miR-142-3p/5p, miR-150, miR-221/ miR-222, miR-363, miR-498 and miR-451 (review summarising findings [86]). The biomarker potential of miRNA has been explored more recently in serum and plasma expression. These studies have tended to focus on established RA and predicting response to therapy [251, 366]. To date there have been no other studies which have studied an at-risk cohort.

5.2 Aims and overview:

In a cohort anti-CCP positive individuals with non-specific MSK symptoms this work aims to:

1) Compare miRNA expression patterns between health and disease along the inflammatory arthritis continuum. This includes individuals at risk of and those that develop RA.
2) Evaluate the change in expression of miRNA from at-risk to early RA.

3) Explore whether miRNAs expression could produce a signature predictive of progression from at-risk to RA.

Two phases were proposed (Figure 17). An initial pilot phase; where miRNAs of interest were identified and a validation phase evaluating the miRNAs of interest in a secondary cohort.

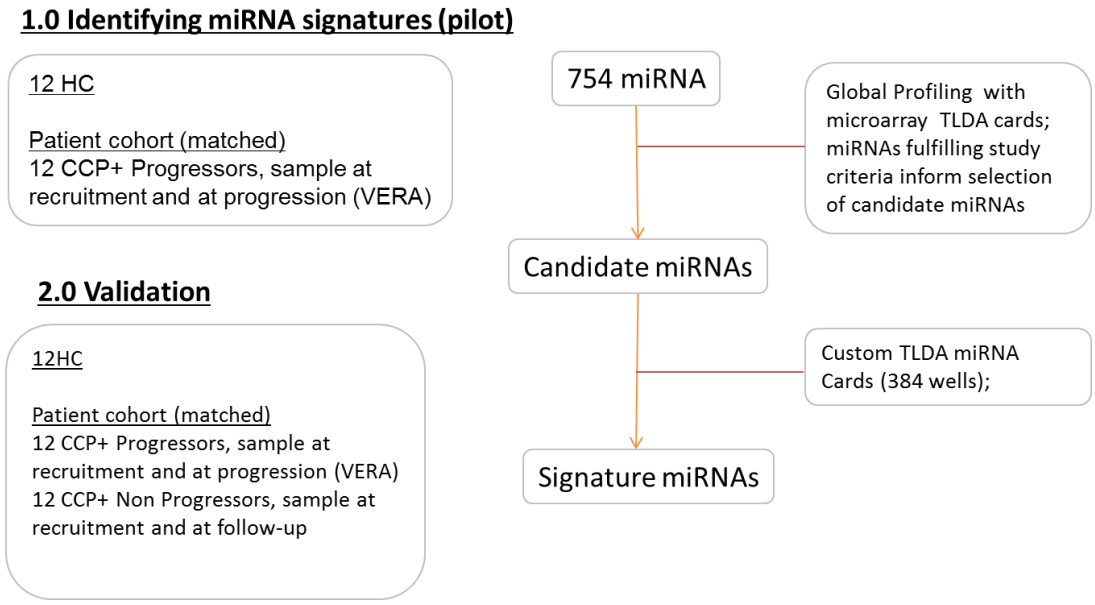


Figure 17 MiRNA study outline.

HC healthy controls; TLDA TaqMan Low Density Arrays; VERA Very early RA.

5.3 Patients and Methods

5.3.1 Patients

Pilot Phase

Previous work has demonstrated ultrasound detectable synovitis at baseline visit in CCP+ individuals [237]. These individuals are in essence considered to already have developed inflammatory pathology, albeit subclinical. For this analysis, therefore only those without ultrasound detectable synovitis (defined as no power Doppler signal) were considered.

Random selection of 12 CCP+ patients who were known to have progressed to RA, termed as very early RA (VERA), were identified (Figure 18A). Blood samples were taken at baseline and at the point of synovitis detection. These samples formed the matched element of the analysis. Twelve healthy controls (HC) were also included. Samples were profiled for >700 miRNAs (5.3.4 Laboratory Method).

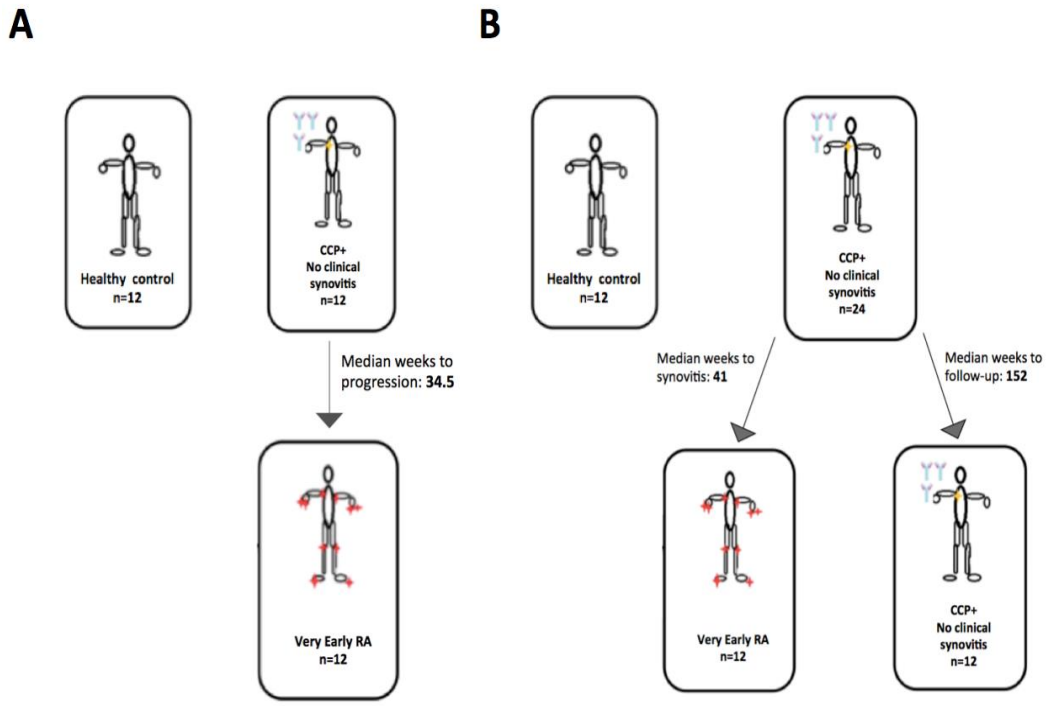


Figure 18 Patients cohorts of serum miRNA profiling.

(17A) *Pilot Phase: Healthy Controls (n=12), CCP positive (CCP) group n=12 progressed to Very Early RA (VERA) group.* ***(17B)*** *Validation Phase including HC n=12, CCP+ progressor and 12 CCP+ non progressor groups.*

Validation Phase

To verify pilot phase findings a further 12 CCP+ progressors were identified. The biomarker potential of candidate miRNAs was evaluated by studying expression in 'negative control' samples. This comparator group consisted of 12 CCP+ individuals who had not progressed to VERA, 'non-progressors' (Figure 18B). It was hypothesised that CCP+ non-progressors would not have the pattern of dysregulation seen in the VERA development of progressors. A further 12 HC were also included. Similarly to the pilot

phase, only those with no evidence of inflammation at baseline ultrasound scan were included.

During the observation period, individuals who progressed to VERA had a matched blood sample at detection of synovitis. The 12 non-progressors had a matched sample 36 weeks from baseline. The 36-week sample was selected following pilot phase data, indicating a median time to progression to synovitis of 34.5 weeks. This enabled closely matched sample point between the two groups. A CCP+ individual was only considered a non-progressor if there had been no VERA development in a minimum of 2 years observation.

Custom cards were used to validate the miRNAs of interest from the pilot phase and the miRNAs previously cited in the literature (Table 17 and 5.3.4 Laboratory Method).

5.3.2 Clinical Assessments

Clinical assessments were performed as previously described (Methods Section 3.5 Protocol). All participants provided (1) baseline demographic details, (2) patient questionnaires, (3) a clinical history of symptoms, (4) MSK examination including a joint count by a rheumatologist (LH & CR). Participants followed the study schedule as outlined in (Chapter 3.5 Study Protocol). Individuals attended 3 monthly visits for the first year and as clinically indicated thereafter, or until they developed inflammatory arthritis; defined by the presence of at least one tender and swollen joint confirmed by a rheumatologist. Blood sampling was performed at baseline and then at regular intervals until the development of IA.

5.3.3 Ultrasound Assessments

Ultrasound examination was performed by a rheumatologist trained in MSK ultrasound (JLN), blinded to the clinical examination [237]. Patients recruited to the study had scans of the wrists, MCPs, PIPs and MTPs bilaterally. The scans were performed at baseline, 6 and 12 months, then annually and at withdrawal to confirm an IA diagnosis. All

participants required no features of power Doppler activity to joints of the hands and feet at baseline assessment.

5.3.4 Laboratory Methods

Isolation and profiling of serum microRNA

Peripheral blood samples were collected in blood clot activator vacutainer tubes. Samples were left at room temperature for the blood to clot and were centrifuged at 3000rpm for 10 minutes. Serum was separated into aliquots, centrifuged at 13,000g for 10 minutes at 4°C. The serum collected was stored at -80°C until further use. Stored serum samples were transferred to Wellcome Trust Brenner Building for microRNA analysis.

Serum microRNAs were isolated according to the manufacturer's protocol using miRNeasy serum plasma kit (Qiagen, UK). The protocol follows the recommendations and incorporates the advised alterations for using RNA purified from serum/plasma (manufactures guidance entitled "Optimised protocols for microRNA profiling with precious samples").

qRT PCR for Global profiling

Isolation of RNA

200 µl of serum was transferred to a clean 1.5 ml microtube. 1 ml QIAzol was added and mixed by pipetting to denature the sample and incubated at room temperature for 10 mins. During incubation, the following was added:

i. 1.25 µl of 0.8 mg/ml MS2 RNA (1 µg in total) (Roche) - RNA stabiliser during cDNA synthesis.
ii. 3.5 µl of 1.6 x 10^8 copies/µl *c.elegans* miR-39 miRNA mimic (Qiagen).

200 µl chloroform was added and shaken vigorously at room temperature for 15 seconds. The protocol supplied with the kit was followed until the final elution, ensuring that the correct volumes, consistent with the starting volume of the serum aliquot, were

used. The final elution was performed with 20 µl molecular grade H_2O. RNA was stored at -80°C until further use.

Complement DNA synthesis (cDNA)

Taqman miRNA reverse transcription kit was used (Life Technologies) for the reverse transcription reaction stage. A reaction volume of 10 µl was recommended including the 3 µl RNA from serum:

RNA input (isolated from serum)	3 µl
10x Megaplex reverse transcription primers*	1 µl
100 mM dNTP	0.27 µl
Multiscribe reverse transcriptase (50 U/µl)	2 µl
10x reverse transcription buffer	1 µl
25 mM $MgCl_2$	1.2 µl
RNase Inhibitor (20 U/µl)	0.13 µl
dH_2O	1.4 µl

* Primer A for Taqman array human miRNA card A and Primer B Taqman array human miRNA card B

The microtubes were placed in a thermal cycler (Life Technologies) set for the following cycling:

16°C 2 minutes
42°C 1 minute } x 40 cycles
50°C 1 second
85°C 5 minutes and then hold at 4°C.

The end reverse transcription product equates to cDNA.

Pre-amplification

For the pre-amplification reaction, 10 µl cDNA from the reverse transcription was used in a reaction volume of 50 µl consisting of:

Reverse transcription product	10 µl
2x Taqman pre-amplification mastermix	25 µl
10x Megaplex pre-amplification primers*	5 µl
dH_2O	10 µl

* Separate reactions required for primer set A and primer set B

The microtubes were placed in a thermal cycler (Life Technologies) set for the following cycling:

95°C 10 minutes

55°C 2 minutes

72°C 2 minutes

95°C 15 seconds
60°C 4 minutes } x 14 cycles

99°C 10 minutes, and then hold at 4°C.

The pre-amplification product was stored at -20°C until it was run on the miRNA Taqman Low density Array (TLDA) array plate.

Preparations for TLDA plate:

The undiluted pre-amplification product was added to the mastermix:

Pre-amplification product	9 µl
2x Taqman Universal Mastermix II, No UNG	450 µl
dH_2O	441 µl

Pre-configured micro fluidic cards were used, enabling the quantification of 754 human miRNAs. Taqman Low density microRNA cards A v3.0 and B set v2.0 (Life Technologies) were prepared according to the instructions in the manuals provided with the Megaplex primers. 100 µl of product per channel was loaded via the fill port using a 100 µL

micropipette. The cards were centrifuge at 1200g 1 minute x2 using Heraeus Multifuge to distribute the reaction mix to the reaction wells. The array cards were sealed using the TaqMan Array Micro Fluidic Card Sealer. Cards were loaded on an Applied Biosystems 7900HT fast real-time system (Life technologies). Data was automatically analysed with SDS Relative Quantification software version 2.4 (Life Technologies).

qRT PCR on custom miRNA cards

TLDA custom cards (Life Technologies) were used in the validation phase once the miRNAs of interest had been established. The TLDA custom cards, comprising of 384 wells, allowed for 31 candidate miRNAs and RNU6B control for normalisation (as per manufacturer recommendation). Extraction of serum RNA was as described in the above methods. Custom primers for the selected miRNAs were used for reverse transcription and pre-amplification steps. Four patient samples could be included on each card; baseline and follow-up samples for pairs of patients were assigned to each card. Data were automatically analysed with SDS Relative Quantification software version 2.4 (Life Technologies).

5.3.5 Statistical Analysis

This project generates large quantities of data despite the relatively small sample numbers, for which a comprehensive analytical approach was therefore required. Formal guidance regarding analysis of micro array data is sparse, however work within the field and a general consensus of opinion has resulted in an accepted approach.

In qRT PCR a positive reaction is detected by accumulation of fluorescent signal. The threshold cycle (Ct) is the fractional cycle number at which the fluorescent signal of the reporter dye passes the threshold (Figure 19). A lower Ct value represents a high level of expression. For this analysis, miRNAs with a Ct <33 were considered as adequately expressed. MiRNAs Ct>33 were therefore excluded, as this would represent weak reaction and possible environmental contamination rather than true expression.

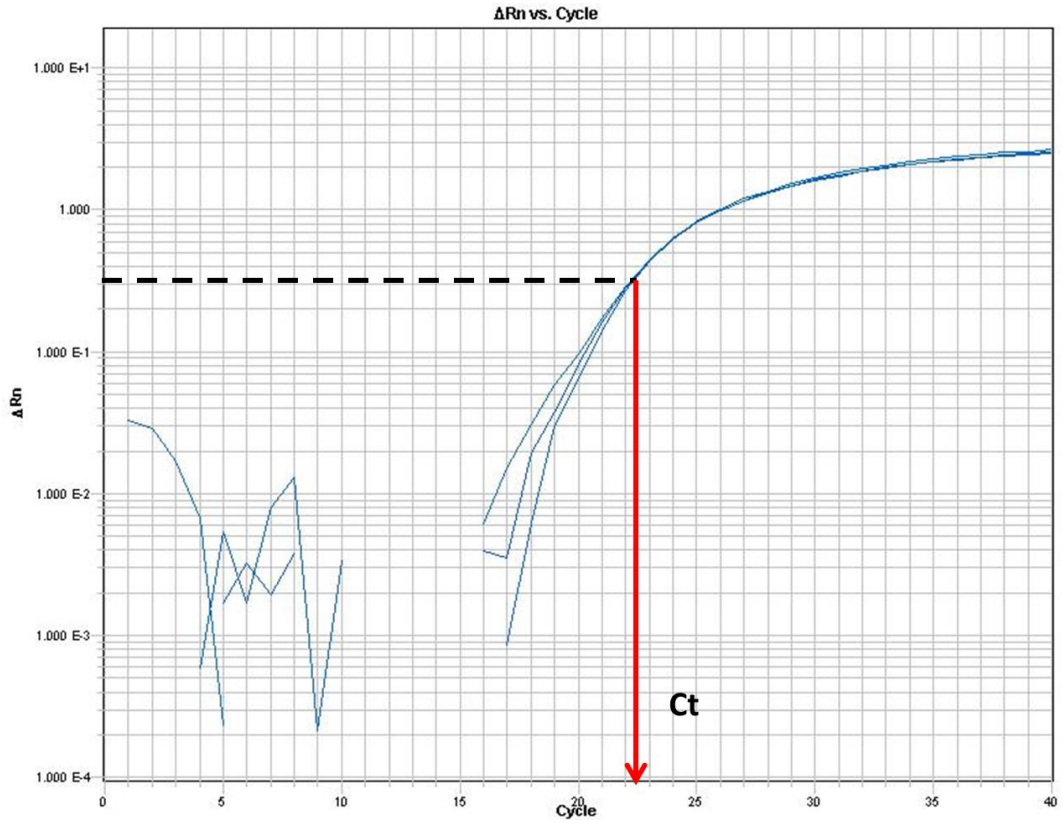


Figure 19 Real-time PCR output of the endogenous control RNU6 on custom cards (triplicate)

The PCR was run for 40 cycles. The point at which the curve intersects the threshold (horizontal dashed black line) is the C_T. The C_T for this miRNA is 22.5.

Pilot phase: Global profiling

There is no current consensus of internal control for plasma miRNA. Previous studies have used specific miRNA (e.g. Let-7a) as endogenous controls, whilst others have used the average value [251]. In this study, normalisation was achieved by calculating the average value across total expressed miRNAs for each sample. Each miRNA was then subtracted from this value to give the delta Ct (dCt). To validate the accuracy of using the average Ct value as a reference for normalization, NormFinder software was used. This software is recommended when assessing the optimum candidate for normalization by determining the most stably expressed microRNA [367]. Reassuringly, the result identified average Ct as the optimal and most stable normalisation method.

To analyse the expression of miRNAs the median and mean dCt of each miRNA across each cohort was calculated. Median was considered most appropriate given the sample size. Mean dCt provided a comparator value for this pilot phase. For 'between-cohort' comparisons, quantile regression, adjusting for age, was used to obtain adjusted between-group differences in median dCt. This was converted into a fold difference (FD). For 'within-patient' changes (CCP+ → VERA matched samples), ddCt was calculated for each within-patient difference, then median fold change (FC) was calculated ($FC= 2^{-[dCt\ (follow\text{-}up)-dCt\ (baseline)]}$). If FD or FC<1, -1/(value) was calculated. Negative values therefore indicate that expression was lower at follow-up compared to baseline (negative FC).

Stringent criteria for the selection of miRNAs of interest was followed for this analysis. The primary criterion stipulated that there should be a FD/FC of four in expression level (FD/FC≥4). Several other miRNA studies have considered FC at lower levels (e.g. FC>2), however a FD/FC≥4 was felt to have greatest biological significance. For 'within-patient' change (CCP+ → VERA, matched samples) pattern of dysregulation had to be consistent across ≥75% of the cohort. Applying these criteria, a list of candidate miRNA was generated.

On an exploratory level, miRNA expression levels were compared using Wilcoxon test to compare paired cohorts; p values <0.05 were considered significant. Given the preliminary nature of the work, descriptive results and inferential statistics are presented.

Validation: MiRNA Custom cards

Custom cards (384 well) allowed for triplicates of each candidate miRNA to be measured. Normalisation was achieved using the mean of the endogenous control replicates (RNU6).

Undetermined Ct values were imputed at the replicate level prior to analysis using R package nondetects, which employs an expectation maximisation algorithm [368]. The algorithm models the probability of an undetermined value occurring as a function of the observed values in the dataset, on the assumption that values may be undetermined due to a failure to amplify, rather than the 'true' value of Ct being >40.

The same analytical approach used in the pilot phase was followed in the validation phase to assess for between group and within patient differences. An additional analysis was required to determine whether specific miRNAs of interest could differentiate at-risk individuals. Baseline expression of candidate miRNAs were compared between the progressor and non-progressor groups to produce a FD. The fold difference was calculated as $FD = 2^{-[dCt\ (progressors)-dCt\ (non\text{-}progressors)]}$. As with FC and FD in the pilot phase, if the value was <1, it was transformed to -1/FD. Negative values therefore indicate that expression was lower in progressors compared to non-progressors (negative FD). The small numbers limits this to an exploratory analysis only.

Associations with clinical variables were assessed using Spearman's rank. Area under the ROC curve for classifying progressors/non-progressors was calculated for each microRNA. Sensitivity and specificity were calculated at the point which maximised the Youden index (sensitivity+specificity-1).

In both phases, raw data was exported from SDS Relative Quantification software to Microsoft excel and then exported to SPSS v.21 and R for statistical analysis. Graphpad Prism 5 was used to produce the figures. Gene Cluster 3.0 and Java TreeView were used to generate MiRNAs heatmaps of the hierarchical clustering analysis.

Power calculations

As this is an exploratory study, no formal power calculations have been performed. For both pilot and validation phases, a *rule of thumb* of n=12 per group was adopted [369].

5.4 Results

5.4.1 Patient Characteristics

At time of analysis, 136 patients that had tested positive for CCP had been recruited to the prospective 'at-risk' clinic. Fifty-seven patients progressed to VERA after a median of 8.6 months (range 0.1-52.4). Of those 57, 29 had no ultrasound-detectable synovitis at baseline of which, 12 were randomly selected for the pilot phase. A further 24 patients (12 that progressed to RA and 12 that did not) were randomly selected for validation phase. Patient characteristics are listed in Table 18.

In the pilot phase, the HC and CCP+ to VERA cohorts were well matched for gender although the HC were younger. In the validation phase there were more females in the progression group compared to the non-progression group. The pilot phase median weeks to progression was 34.5 (IQR 13.5-56.5) compared to 41 (IQR 25.8-65.0) weeks in the validation phase. Individuals within the non-progression cohort were followed up for a minimum period of 127 weeks.

Table 18 Baseline characteristics of individuals for pilot and validation phases

Pilot phase	HC	CCP+ progression to VERA	
Number	12	12	
Median Age (IQR)	43 (38-55.3)	52 (43-70)	
Female	8 (67%)	8 (67%)	
Median weeks to synovitis (IQR)	-	34.5 (13.5-56.5)	
Validation phase	**HC**	**CCP+ progression to VERA**	**CCP+ non-progression**
Number	12	12	12
Median Age (IQR)	35.5 (29.8-43.8)	54 (39.0-59.2)	53.5 (50.0-65.5)
Female	6 (50%)	10 (83.3)	6 (50)

Median weeks to synovitis (IQR)	-	41 (25.8-65.0)	-
Median weeks follow-up (IQR)	-	-	152 (127.3-241.5)

(HC Healthy controls, IQR Interquartile range, VERA Very early rheumatoid arthritis)

5.4.2 Pilot phase

5.4.2.1 Comparing health with CCP+ and VERA cohorts:

Table 19 lists the miRNAs that were dysregulated between the cohorts. For between cohort comparisons, miRNAs included in the table had a FD >4. For matched samples miRNAs with FC>4 and consistent pattern of dysregulation were selected. Between HC group (n=12) and CCP+ (n=12), 8 dysregulated miRNAs were identified (4 down- and 4 up-regulated); between HC and (matched CCP+ to) VERA cohort (n=12), 13 dysregulated miRNAs were recorded (to note, all bar one were upregulated). The greatest FD was seen with miR-628-5p comparing health with CCP and VERA (FD 1176.3 and 59.7 respectively). These FDs were significantly higher than the remainder of the reported FDs.

In addition to the 17 miRNA in Table 19 that met the criteria, there were several other miRNAs that demonstrated substantial changes. MiR-579 was identified to be of interest as the mean FC 4.27 (despite the median value FC being <4). Furthermore, this miRNA had consistent dysregulation in 11/12 patients. In order to be inclusive of potential signals that may become significant with a powered study, this miRNA was highlighted. For miR-15b and -335 the Ct was >32 for some CCP+ patients at baseline. As a result, the calculated fold changes may not be accurate. These miRNAs were not deemed to have fulfilled the criteria for dysregulation but were retained for further investigation because the calculated FC was near or above the design cut-off. On reviewing the raw data for these miRNAs, a pattern of expression emerged from CCP to VERA despite not having a significant fold change. Furthermore, in the matched analysis all progressors showed consistent dysregulation.

		HC	CCP	VERA	CCP vs. HC	VERA vs. HC	CCP to VERA (within progressors)		
	miR	dCt median (IQR)	dCt median (IQR)	dCt median (IQR)	FD between medians	FD between medians	Median (IQR) ddCt	Median FC	N upregulated (/12)
1	miR-16	-6.3 (-7.1, -6.0)	-7.1 (-7.6, -7.0)	-7.6 (-8.2, -7.4)	1.7	2.5	-0.4 (-1.1, -0.1)	1.3	10
2	miR-18a	0.1 (-0.8, 1.3)	1.3 (0.3, 2.0)	-0.1 (-1.0, 0.4)	-2.3	1.1	-1.6 (-2.0, -1.0)	3.1	10
3	miR-19a	-0.9 (-1.6, 1.0)	-2.4 (-2.8, -1.9)	-3.2 (-3.5, -2.9)	2.8	4.9↑	-0.6 (-1.7, 0.1)	1.5	9
4	miR-21	-2.8 (-4.1, -2.5)	-3.8 (-4.4, -3.3)	-4.3 (-4.8, -4.1)	2.0	2.8	-0.7 (-1.2, -0.2)	1.6	9
5	miR-22	4.2 (0.5, 4.8)	3.0 (1.5, 5.1)	0.9 (0.6, 1.5)	2.3	9.8↑	-2.1 (-3.6, -1.5)	**4.3↑**	**12**
6	miR-26b	-1.4 (-3.0, 0.2)	-3.1 (-3.4, -1.8)	-3.6 (-4.2, -2.8)	3.2	4.6↑	-0.7 (-2.3, -0.3)	1.7	10
7	miR-34a	-0.2 (-2.1, 0.7)	-0.1 (-0.3, 1.0)	-0.6 (-2.1, 0.2)	-1.1	1.3	-0.1 (-0.9, 0.3)	1.1	6
8	miR-101	2.5 (1.7, 3.2)	1.6 (1.3, 1.8)	0.4 (-0.3, 0.9)	1.9	4.3↑	-1.1 (-1.9, -0.6)	2.1	11
9	miR-132	-1.6 (-2.0, -1.4)	-1.7 (-2.1, -1.6)	-2.5 (-2.7, -2.2)	1.1	1.9	-0.8 (-1.1, -0.3)	1.7	11
10	miR-142-3p	-2.3 (-4.5, -1.3)	-4.4 (-4.5, -3.8)	-5.0 (-5.2, -4.5)	4.3↑	6.5↑	-0.4 (-1.1, -0.3)	1.4	10
11	miR-142-5p	3.9 (2.9, 5.7)	3.0 (2.4, 5.5)	1.7 (1.5, 4.3)	1.9	4.6↑	-1.3 (-1.4, -0.3)	2.4	10
12	miR-146a	-7.3 (-7.5, -6.5)	-7.1 (-7.6, -7.0)	-7.5 (-7.7, -7.3)	-1.1	1.1	-0.5 (-0.8, 0.2)	1.4	8
13	miR-155	-0.9 (-2.3, 1.4)	0.2 (-0.5, 1.1)	-0.5 (-0.7, -0.3)	-2.1	-1.3	-0.3 (-1.3, 0.1)	1.2	8
14	miR-195	-2.5 (-2.9, -1.6)	-0.4 (-3.4, 0.2)	-2.9 (-4.2, -0.5)	-4.3↓	1.3	-1.1 (-2.0, -0.5)	2.1	11
15	miR-197	-3.2 (-2.8, -1.3)	0.6 (-1.5, 2.9)	0.6 (-2.5, 2.2)	-13.9↓	-13.9↓	-0.6 (-2.7, 1.3)	1.5	7
16	miR-203	2.3 (1.8, 3.0)	3.0 (2.1, 3.0)	2.9 (2.6, 3.4)	-1.6	-1.5	0.0 (-0.4, 0.6)	1.0	6
17	miR-210	3.9 (0.8, 5.0)	3.1 (2.2, 3.6)	1.5 (1.1, 1.8)	1.7	5.3↑	-1.8 (-2.9, -0.3)	3.4	10
18	miR-223	-9.3 (-9.6, -8.7)	-9.8(-9.9, -9.5)	-10.3 (-10.5, -10.2)	1.4	2.0	-0.4 (-0.6, 0.2)	1.3	6/9*
19	miR-361	-0.1 (-1.1, 1.1)	2.1 (-0.5, 2.5)	0.1 (-0.2, 0.7)	-4.6↓	-1.1	-1.5 (-2.1, -1.0)	2.9	11
20	miR-374‡	-2.2 (-3.2, -2.1)	-3.2 (-3.7, -2.0)	-4.1 (-4.3, -3.9)	2.0	3.7	-0.6 (-1.0, -0.4)	1.5	12

21	miR-382	-0.7 (-1.2, 0.5)	1.4 (0.9, 2.7)	-0.1 (-1.2, 0.5)	-4.3↓	-1.5	-2.0 (-2.8, -0.8)	**4.1↑**	**11**
22	miR-454‡	-1.3 (-2.8, 0.1)	-2.3 (-3.2, -0.1)	-3.0 (-3.3, -1.3)	2.0	3.2	-0.5 (-1.1, -0.2)	1.4	10
23	miR-486-3p	4.3 (2.5, 5.6)	4.9 (2.5, 6.2)	3.5 (3.1, 4.3)	-1.5	1.7	-2.0 (-3.1, 0.2)	**4.1↑**	**9**
24	miR-520c-3p	2.3 (0.3, 2.9)	-0.4 (-1.6, 2.4)	-1.1 (-1.4, 0.5)	6.5↑	10.6↑	0.3 (-1.7, 0.6)	-1.3	5
25	miR-579§	4.0 (3.5, 5.0)	3.5 (3.1, 3.7)	2.8 (1.8, 2.9)	1.4	2.3	-1.2 (-1.7, -0.3)	2.2	11
26	miR-590-3P	5.2 (2.9, 5.8)	3.2 (2.3, 4.2)	2.4 (2.2, 3.1)	4.0↑	7.0↑	-0.9 (-1.1, 0.2)	1.9	8
27	miR-590-5p	1.2 (0.7, 1.4)	-0.4 (-1.2, -0.2)	-1.3 (-1.6, -0.7)	3.0	5.7↑	-0.5 (-1.1, -0.3)	1.4	10
28	miR-598	3.7 (3.1, 4.3)	2.8 (2.5, 3.6)	1.5 (1.4, 1.9)	1.9	4.6↑	-1.2 (-1.7, -0.8)	2.2	12
29	miR-628-5p	3.0 (-7.5, 7.9)	-7.2 (-8.3, 3.9)	-2.9 (-7.8, 2.2)	1176.3↑	59.7↑	0.4 (-1.6, 1.7)	-1.3	5
30	miR-15b#	10.1 (3.7, 11.9)	6.5 (5.6, 8.8)	3.5 (1.9, 4.7)	12.1†	97.0†	-2.6 (-4.0, -1.8)	6.1†	11/11
31	miR-335#	6.1 (5.6, 8.3)	6.5 (5.2, 8.1)	4.2 (3.9, 5.1)	-1.3†	3.7†	-1.9 (-3.9, -1.0)	3.8†	12

Table 19 List of miRNAs of interest with age-adjusted FC≥4 between the three studied cohorts

(↑upregulated FC≥4, ↓downregulated FC≤-4). MiRNA highlighted in bold in matched samples (CCP-VERA) met criteria of median FC≥4 and ≥75% consistent dysregulation. MiRNA previously reported in the literature are highlighted by shading. FD=fold difference; FC=fold change. If FD<1, FD=-1/FD. Estimates for each cohort were obtained at the mean age (52 years).

**For 3 patients, Ct values at follow-up were extremely low (all ≈2, compared to ≈14 for the rest); these 3 values were considered to be inaccurate and in a conservative approach were excluded from analysis.*

†In these miRNAs, Ct was >32 for some healthy controls & CCP+ patients at baseline. As a result, the calculated fold differences & changes may not be accurate; therefore, these genes were not deemed to have fulfilled our criteria for dysregulation but were retained for further investigation in the validation cohort because the calculated FC were near or above our cut-off and all progressors showed consistent dysregulation.

§In this MiRNA mean FC>4; miR-579 FC 4.27 (although median FC<4), furthermore consistent dysregulation seen in 11/12 patients. As custom cards had capacity for 31 miRNAs to be evaluated these two miRNAs were selected as potentially important.

‡The original selection of miR-454 (FD 4.1 between HC-VERA) and miR-374 (FD 4.6 between HC-VERA) from the pilot phase was based on unadjusted between-group differences; following age-adjustment they no longer met criteria.

5.4.2.2 CCP+ to VERA matched sera miRNA evaluation:

From paired analysis of the matched samples, the within-patient change of 3 circulating miRNAs were upregulated upon progression from CCP+ state to VERA (see Table 19, highlighted in bold, end column). Serum miR-22 expression increased the most in patients from CCP+ to VERA [median FC 4.3 (IQR 2.8, 12.1; *p=0.005*)]. The expression of this miRNA increased in all patients (12/12). Comparable levels of upregulation were found for miR-382 [median FC 4.1 (IQR 1.7, 6.9; *p=0.002*; increased in 11/12] and miR-486-3p [median FC 4.1 (IQR 0.9, 8.6; *p=0.027*; increased in 9/12]. Figure 20 represents the change in expression from CCP+ to VERA of these 3 miRNA.

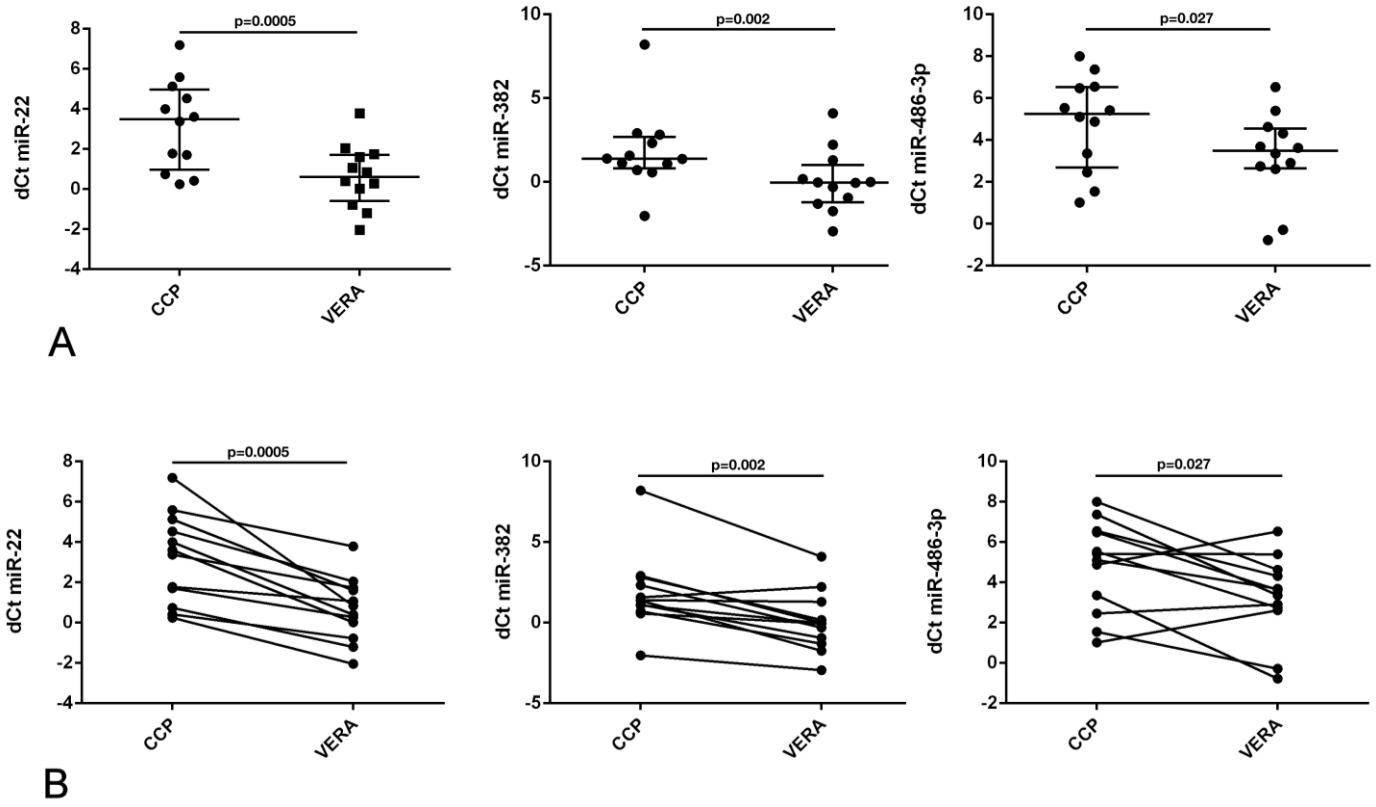


Figure 20 Comparison of expression levels of miR-22, miR-382 and miR-486-3p in matched samples from CCP+ to VERA

(A) Dot plot with median dCt represented by the middle horizontal line and 1st-3rd IQR represented by the whiskers. (B) represent median dCt from CCP+ to VERA per matched sample. A lower dCt value represents a higher level of expression.

MiR-628-5p, which had demonstrated very high levels of dysregulation in the CCP and VERA cohorts compared to health, did not meet the criteria in the paired analysis with a FC of -1.3. This may be attributed to the majority of change in expression occurring in transition from health to autoimmunity and hence the transition to VERA is minimal. MiR-15b had a FC of 6.1, however it was noted that this miRNA had great variability in the Ct values and must be reviewed cautiously.

5.4.2.3 Cited MiRNA associated with RA pathogenesis:

The shaded rows featured in Table 19 indicate the 9 miRNAs frequently cited to have associations with RA. None of these miRNAs reached the stringent FC/FD criteria set in this study. Between cohort comparisons, miR18a had the highest FD of -2.3 between CCP+ and health. Within the matched samples this miRNA had the highest FC 3.1 with a consistent pattern of dysregulation; upregulated in 10/12 individuals. MiR-146a and miR-155 are the two most cited miRNAs in relation to RA. Within this set of samples, both miRNA were upregulated in the majority (8/12) of the individuals who progressed to RA (Figure 21 A & B); however, the FCs were not substantive (median 1.4 and 1.2 respectively).

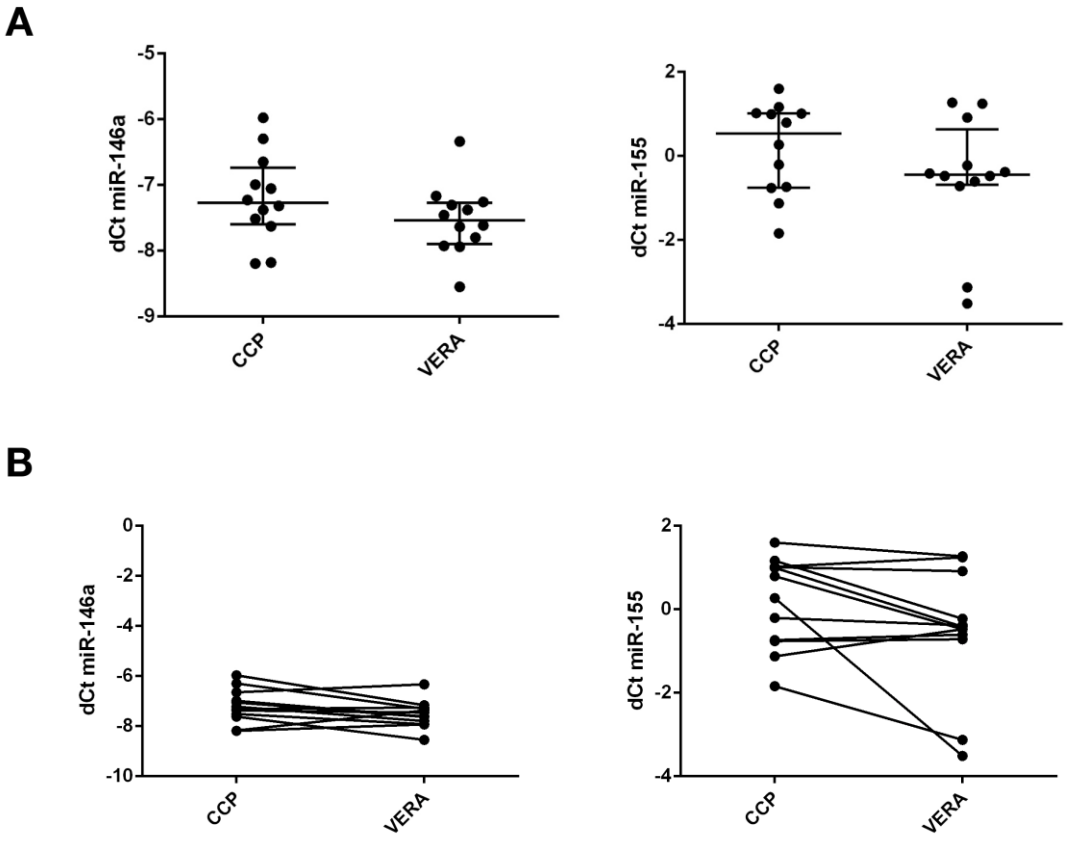


Figure 21 Serum miR146a and miR-155 expression in the pilot phase.

(A) Relative expression levels of miR-146a and miR-155 in CCP and matched VERA groups, dot plot with median dCt represented by the middle horizontal line and 1st-3rd IQR represented by the whiskers (B) Comparison of expression levels (median dCt) of miR-146a and miR-155 in the matched samples: miR-146a FC 1.4 (-1.2, 1.8) and miR-155 FC 1.2 (-1.1, 2.6) (dCT, delta threshold cycle).

The dysregulated MiRNAs highlighted from the between cohort comparisons and the 3 miRNA identified in the matched analysis formed the basis of the tailored array cards. In addition, the 9 cited miRNAs involved in the pathogenesis in RA were included. In total, 31 miRNAs were evaluated in the validation phase, maximising the use of the 384 well custom cards.

5.4.2.4 Hierarchical clustering

In order to ascertain whether expression of specific miRNAs grouped or clustered together within the cohorts, an unsupervised hierarchical clustering analysis was performed. Using complete linkage, cluster analysis of the global expression profiles of the 31 miRNAs was generated Figure 22.

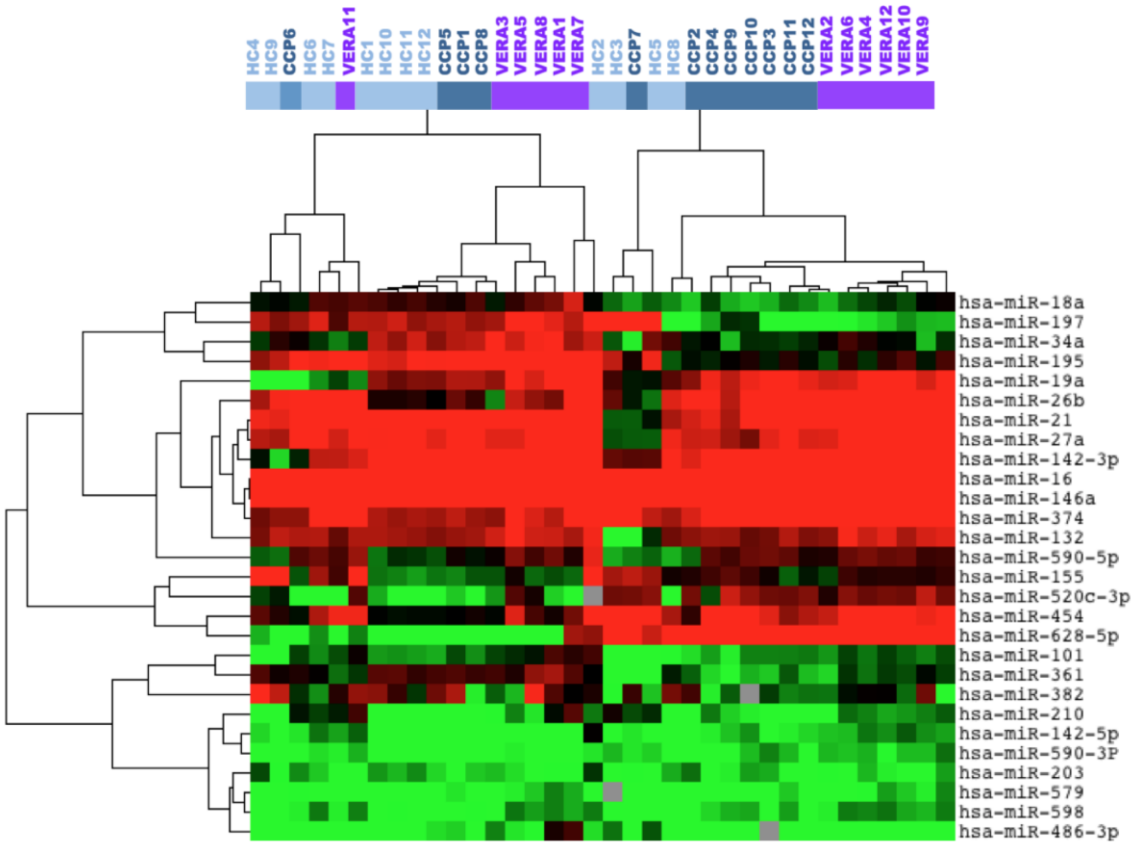


Figure 22 MiRNAs heatmaps were generated using hierarchical clustering

Green indicates low expression; Red indicates high expression levels. (Generated using software Gene Cluster 3.0 and Java TreeView). Top bar identifies each sample; light blue = HC samples, dark blue = CCP+ samples, purple = VERA samples.

While the cohorts do not completely cluster into 3 distinct groups there was clear clustering of several samples within each cohort. This highlights clear similarities in patterns of expression between CCP+ and VERA. There was also greater variation in the healthy control samples.

5.4.3 Validation phase

A list of the dysregulated miRNAs from the three studied cohorts (22 in total) and 9 miRNAs (miR-21, miR-146a, miR-155, miR-18a, miR-34a, miR-203, miR-223, miR-16, and miR-132) that had been cited in the literature was thus established.

To validate the findings, a further 24 CCP+ patients were randomly selected. This consisted of 12 CCP+ patients who progressed to VERA (progressor cohort) and had a matched blood sample available at detection of synovitis. A comparator group, consisting of 12 CCP+ individuals that did not progress to VERA (non-progressor cohort) were also selected. (Figure 18 and Table 18). The 31 candidate miRNAs were then quantified.

5.4.3.1 Comparing Health with CCP+ (progressor and non progressor) and VERA cohorts:

Expression profiles of RNU6B (manufacturers recommended control) were stable across all samples with Ct values ranging between 23 and 26. Table 20 lists the miRNAs that were dysregulated between the cohorts, the arrows indicate where there has been consistent dysregulation (FD >4) in both pilot and validation phase. In contrast to the pilot phase, only 3 miRNAs (miR-22, -34a, 335#) were dysregulated (all up-regulated) between validation HC group and validation CCP+ progressors. MiR-22 demonstrated significant upregulation with a FD 11.3. The inclusion of the CCP+ non-progressor cohort provided an additional comparison to health. Only miR-590-3P met the criterion of dysregulation compared to health. Between HC group and VERA cohort, 6 dysregulated miRNAs were identified (all upregulated); and 4 miRNAs were validated from the pilot results (as indicated by (↑)); miR-19a, miR-22, miR-590-3p and miR-598. MiR-628-5p high expression in the pilot phase failed to be validated in this phase.

Table 20 List of miRNAs of interest with age-adjusted FD≥4 between the three studied cohorts in the validation phase.

	HC	CCP (non-prog)	CCP (prog)	VERA	CCP (non-prog) vs. HC	CCP (prog) vs. HC	VERA vs. HC
miR	dCt median (IQR)	dCt median (IQR)	dCt median (IQR)	dCt median (IQR)	FD between medians	FD between medians	FD between medians
miR-16	-7.9 (-8.8, -6.8)	-8.2 (-8.8, -7.2)	-8.9 (-10.5, -7.8)	-9.5 (-11.3, -8.0)	1.2	2.0	3.0
miR-18a	0.1 (-0.6, 0.6)	-1.1 (-1.3, -0.7)	-1.1 (-2.8, -0.5)	-1.9 (-3.0, -0.4)	2.3	2.3	4.0
miR-19a	-0.8 (-0.9, -0.4)	-1.7 (-2.3, -1.4)	-1.9 (-3.3, -1.0)	-2.8 (-3.6, -1.2)	1.9	2.1	4.0 (↑)
miR-21	-2.1 (-2.4, -1.6)	-2.6 (-3.4, -2.4)	-3.7 (-4.9, -3.0)	-4.3 (-5.3, -1.5)	1.4	3.0	4.6
miR-22	**6.6 (3.8, 9.0)**	**7.4 (4.1, 8.2)**	**3.1 (1.8, 7.3)**	**3.3 (1.0, 4.3)**	**-1.7**	**11.3**	**9.8 (↑)**
miR-26b	-0.4 (-1.0, -0.2)	-1.5 (-1.9, -1.0)	-1.3 (-3.2, -0.7)	-1.8 (-3.4, -0.7)	2.1	1.9	2.6
miR-34a	3.4 (2.4, 4.8)	3.0 (1.7, 4.0)	1.4 (0.7, 3.2)	2.5 (1.0, 3.1)	1.3	4.0	1.9
miR-101	2.3 (2.3, 3.7)	3.2 (2.6, 3.7)	2.1 (0.5, 3.3)	1.6 (0.2, 3.3)	-1.9	1.1	1.6
miR-132	0.8 (-0.2, 1.4)	0.2 (-0.6, 1.3)	-0.2 (-1.9, 0.1)	-0.7 (-1.7, 0.3)	1.5	2.0	2.8
miR-142-3p	-3.8 (-3.7, -3.0)	-3.9 (-5.1, -3.5)	-4.4 (-5.2, -3.7)	-4.4 (-5.2, -2.6)	1.1	1.5	1.5
miR-142-5p	2.0 (1.9, 2.5)	1.7 (0.8, 1.9)	1.4 (-0.1, 2.3)	0.6 (-0.4, 2.5)	1.2	1.5	2.6
miR-146a	-6.2 (-6.9, -4.7)	-6.1 (-7.3, -5.3)	-7.3 (-8.3, -6.4)	-7.1 (-8.3, -6.6)	-1.1	2.1	1.9
miR-155	-0.4 (-1.0, 0.4)	-0.8 (-1.7, 0.3)	-1.3 (-1.8, -0.4)	-1.3 (-1.9, -1.1)	1.3	1.9	1.9
miR-195	-1.4 (-2.0, -0.9)	-1.9 (-2.4, -0.7)	-2.9 (-3.6, -1.4)	-3.0 (-3.5, -1.7)	1.4	2.8	3.0
miR-197	-2.2 (-2.8, -1.1)	-2.6 (-3.6, -1.7)	-4.0 (-4.2, -2.5)	-3.3 (-3.9, -2.9)	1.3	3.5	2.1
miR-203	5.3 (4.8, 7.8)	6.0 (4.6, 7.6)	5.2 (3.5, 6.0)	5.2 (3.9, 5.9)	-1.6	1.1	1.1
miR-210	1.9 (1.3, 2.6)	0.9 (0.3, 1.7)	1.3 (-0.5, 1.7)	0.1 (-0.7, 2.2)	2.0	1.5	3.5
miR-223	-10.7 (-11.1, -10.4)	-10.7 (-12.0, -9.9)	-11.5 (-12.4, -11.1)	-12.2 (-13.2, -10.9)	1.0	1.7	2.8
miR-361	2.4 (1.6, 3.3)	3.3 (1.8, 3.5)	1.6 (0.5, 2.5)	1.7 (0.5, 2.3)	-1.9	1.7	1.6
miR-374	0.1 (-0.3, 0.3)	-0.7 (-1.3, 0.1)	-0.8 (-1.3, -0.4)	-0.8 (-2.5, 0.3)	1.7	1.9	1.9
miR-382	**1.3 (0.9, 2.2)**	**1.1 (0.0, 1.8)**	**-0.2 (-0.5, 1.9)**	**0.5 (-0.2, 1.0)**	**1.1**	**2.8**	**1.7**

miR-454	-1.3 (-1.6, -0.7)	-2.0 (-2.6, -1.7)	-2.0 (-2.6, -1.1)	-2.2 (-3.1, -1.1)	1.6	1.6	1.9
miR-486-3p	**4.0 (2.8, 5.1)**	**3.4 (1.7, 3.9)**	**3.9 (2.6, 5.0)**	**2.6 (1.9, 4.0)**	**1.5**	**1.1**	**2.6**
miR-520c-3p	-2.5 (-3.3, -0.8)	-2.2 (-4.1, -2.1)	-2.9 (-4.8, -2.7)	-3.1 (-4.4, -2.3)	-1.2	1.3	1.5
miR-579	5.1 (4.2, 5.9)	5.4 (4.4, 6.1)	3.9 (2.2, 5.9)	3.9 (2.4, 4.7)	-1.2	2.3	2.3
miR-590-3P	7.9 (6.7, 7.4)	5.4 (4.1, 7.7)	6.3 (5.1, 8.4)	5.9 (5.4, 7.7)	5.7	3.0	4.0 (🡅)
miR-590-5p	2.8 (2.7, 3.2)	2.6 (1.7, 3.8)	1.9 (1.3, 3.1)	2.3 (0.7, 4.3)	1.1	1.9	1.4
miR-598	3.2 (2.2, 3.7)	2.1 (1.4, 3.3)	1.6 (0.9, 2.2)	1.1 (0.0, 2.7)	2.1	3.0	4.3 (🡅)
miR-628-5p	4.1 (3.3, 5.3)	3.4 (2.4, 3.6)	3.5 (2.7, 4.4)	2.6 (2.1, 3.3)	1.6	1.5	2.8
miR-15b#	1.5 (1.0, 1.5)	0.7 (0.4, 1.3)	0.5 (-0.1, 1.6)	0.2 (-1.0, 0.6)	1.7	2.0	2.5
miR-335#	5.3 (4.2, 6.6)	4.2 (2.9, 5.4)	2.8 (2.2, 3.7)	2.2 (1.6, 4.4)	2.1	5.7	8.6

🡅upregulated FC≥4 in both phases, 🡇downregulated FC≤-4 in both phase. MiRNA highlighted in bold in matched samples (CCP-VERA) met criteria of median FC≥4 and ≥75% consistent dysregulation in the pilot phase. FD=fold difference; FC=fold change. If FD<1, FD=-1/FD. Estimates for each cohort were obtained at the mean age (48 years).

5.4.3.2 Baseline miRNA profile of CCP+ progressor compared to CCP+ non-progressor:

Descriptively, there was varied expression of the selected miRNAs at baseline between the progressors and non-progressors. Overall, baseline readings of miRNAs were mostly upregulated in the progressors, compared to the non-progressors. There were 5 miRNAs which were downregulated; miR-26b, miR-210, miR-486-3p, miR-590-3p and miR-628-5p (Table 21). Baseline FD between the two groups for the 3 miRNAs of interest, miR-22, miR-382 and miR-486-3p were 19.7, 2.5 and -1.4 respectively. At baseline, miR-22 was only expressed at Ct<32 in 6/12 non-progressors compared to 10/12 progressors, with a high fold difference between groups (FD 19.7). This supports supporting pilot findings that miR-22 is possibly associated with progression to VERA.

Table 21 Comparison of miRNA expression between baseline samples within the CCP+ progressors and CCP+ non-progressors cohorts.

miRNA	CCP+ non-progressors	CCP+ progressors	Progressors vs. non-progressors
	B/L median dCt (IQR)	B/L median dCt (IQR)	FD between medians
miR-16	-8.2 (-8.8, -7.2)	-8.9 (-10.5, -7.8)	1.6
miR-18a	-1.1 (-1.3, -0.7)	-1.1 (-2.8, -0.5)	1.0
miR-19a	-1.7 (-2.3, -1.4)	-1.9 (-3.3, -1.0)	1.1
miR-21	-2.6 (-3.4, -2.4)	-3.7 (-4.9, -3.0)	2.1
miR-22	**7.4 (4.1, 8.2)**	**3.1 (1.8, 7.3)**	**19.7†**
miR-26b	-1.5 (-1.9, -1.0)	-1.3 (-3.2, -0.7)	-1.1
miR-34a	3.0 (1.7, 4.0)	1.4 (0.7, 3.2)	3.0
miR-101	3.2 (2.6, 3.7)	2.1 (0.5, 3.3)	2.1
miR-132	0.2 (-0.6, 1.3)	-0.2 (-1.9, 0.1)	1.3
miR-142-3p	-3.9 (-5.1, -3.5)	-4.4 (-5.2, -3.7)	1.4
miR-142-5p	1.7 (0.8, 1.9)	1.4 (-0.1, 2.3)	1.2
miR-146a	-6.1 (-7.3, -5.3)	-7.3 (-8.3, -6.4)	2.3
miR-155	-0.8 (-1.7, 0.3)	-1.3 (-1.8, -0.4)	1.4
miR-195	-1.9 (-2.4, -0.7)	-2.9 (-3.6, -1.4)	2.0
miR-197	-2.6 (-3.6, -1.7)	-4.0 (-4.2, -2.5)	2.6

miR-203	6.0 (4.6, 7.6)	5.2 (3.5, 6.0)	1.7†
miR-210	0.9 (0.3, 1.7)	1.3 (-0.5, 1.7)	-1.3
miR-223	-10.7 (-12.0, -9.9)	-11.5 (-12.4, -11.1)	1.7
miR-361	3.3 (1.8, 3.5)	1.6 (0.5, 2.5)	3.2
miR-374	-0.7 (-1.3, 0.1)	-0.8 (-1.3, -0.4)	1.1
miR-382	**1.1 (0.0, 1.8)**	**-0.2 (-0.5, 1.9)**	**2.5**
miR-454	-2.0 (-2.6, -1.7)	-2.0 (-2.6, -1.1)	1.0
miR-486-3p	**3.4 (1.7, 3.9)**	**3.9 (2.6, 5.0)**	**-1.4**
miR-520c-3p	-2.2 (-4.1, -2.1)	-2.9 (-4.8, -2.7)	1.6
miR-579	5.4 (4.4, 6.1)	3.9 (2.2, 5.9)	2.8
miR-590-3P	5.4 (4.1, 7.7)	6.3 (5.1, 8.4)	-1.9†
miR-590-5p	2.6 (1.7, 3.8)	1.9 (1.3, 3.1)	1.6
miR-598	2.1 (1.4, 3.3)	1.6 (0.9, 2.2)	1.4
miR-628-5p	3.4 (2.4, 3.6)	3.5 (2.7, 4.4)	-1.1†
miR-15b#*	0.7 (0.4, 1.3)	0.5 (-0.1, 1.6)	1.1
miR-335#*	4.2 (2.9, 5.4)	2.8 (2.2, 3.7)	2.6†

†In these miRs, Ct was >32 for some patients and/or some values were undetermined and had to be imputed. MiRNAs highlighted in bold satisfied criteria for dysregulation in the pilot phase.

5.4.3.3 Change in miRNA expression within matched sera of CCP+ to VERA progressor and CCP+ to non-progressor:

In the validation phase, no miRNA reached the pre-defined criteria of FC≥4 with consistent dysregulation across ≥75% of the cohort. From the three key miRNAs identified in the pilot phase, miR-486-3p increased in progressors by a median (IQR) FC 2.2 (0.4, 6.0) with upregulation in 7/12 (Table 22,

Figure 23 (A)) compared to stable expression within the non-progressor cohort FC 1.0 (0.7, 3.0). Despite miR-22 demonstrating a significant FD (19.7) between baseline samples of progressors versus non-progressors, the median FC (IQR) within patients was increased in both groups. This was found to be greatest within the non-progressors, FC 3.4 (0.5, 12.1) versus FC 2.5 (0.5, 19.7) in progressors (Table 22,

Figure 23 (B)). Furthermore, it was upregulated in all 12 of the non-progressors compared to 8/12 progressors. This result is not consistent with findings to date but

may reflect miR-22's role in CCP+ individuals with higher expression in individuals that develop IA. Similarly, with miR-382, a greater median FC (IQR) was observed in the non-progressor cohort [FC 2.4 (1.0, 2.6)] versus progressor [FC 1.2 (0.5, 2.6)] (

Figure 23 (A)). This may reflect association of this miRNAs with the CCP+ phenotype.

Two miRNAs were upregulated in the non-progressors baseline to follow-up but not in the progressor group: miR-203 [FC 3.1 (0.5, 6.9) vs. -1.1 (0.3, 4.3)] and miR-579 [FC 3.2 (1.0, 3.7) vs. -1.1 (0.3, 3.3)] (Table 22,

Figure 23 (B)). For both miRNAs, the pattern of dysregulation in the non progressor cohort was consistent in 9/12 of the individuals. MiR-203 has previously been identified as a miRNA involved in RA [87], but has not been studied before in CCP+ at-risk individuals. Whilst, miR-579 was upregulated in VERA compared to HC; the significance of upregulation in both the inflammatory cohorts and the CCP+ non-progressors remains unclear.

MiR-15b and -335 were included in the selected 31 miRNAs following calculated FC which were near or above the criteria of the pilot phase. However, in the validation phase expression in both CCP progressors and non-progressors was stable.

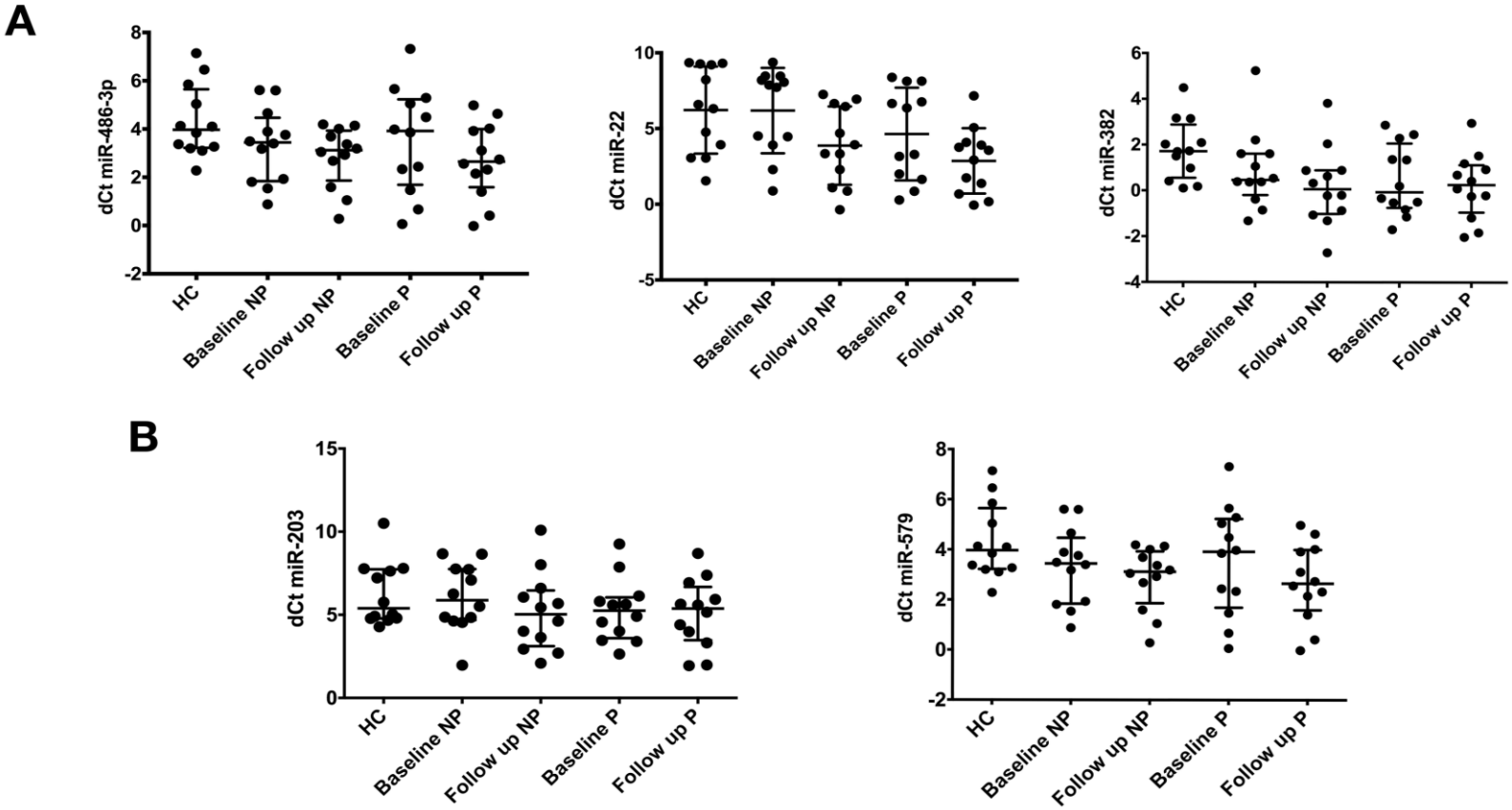


Figure 23 Expression levels of candidate miRNA in serum of individuals from validation phase.

*Baseline and follow-up relative expression in the progressor (P) and non-progressor (NP) cohorts of **(A)** miR-486-3p, miR-22 and miR-382 **(B)** miR-203 and miR-579.*

5.4.3.4 Hierarchical clustering

An unsupervised hierarchical clustering analysis of the global expression profiles of 31 miRNAs of interest from the validation phase was generated using complete linkage. The cohorts analysed replicated those presented in the pilot phase and therefore the CCP+ non-progressor group was excluded. As with the pilot phase findings, there are patterns of similarity with clustering of individuals in cohorts.

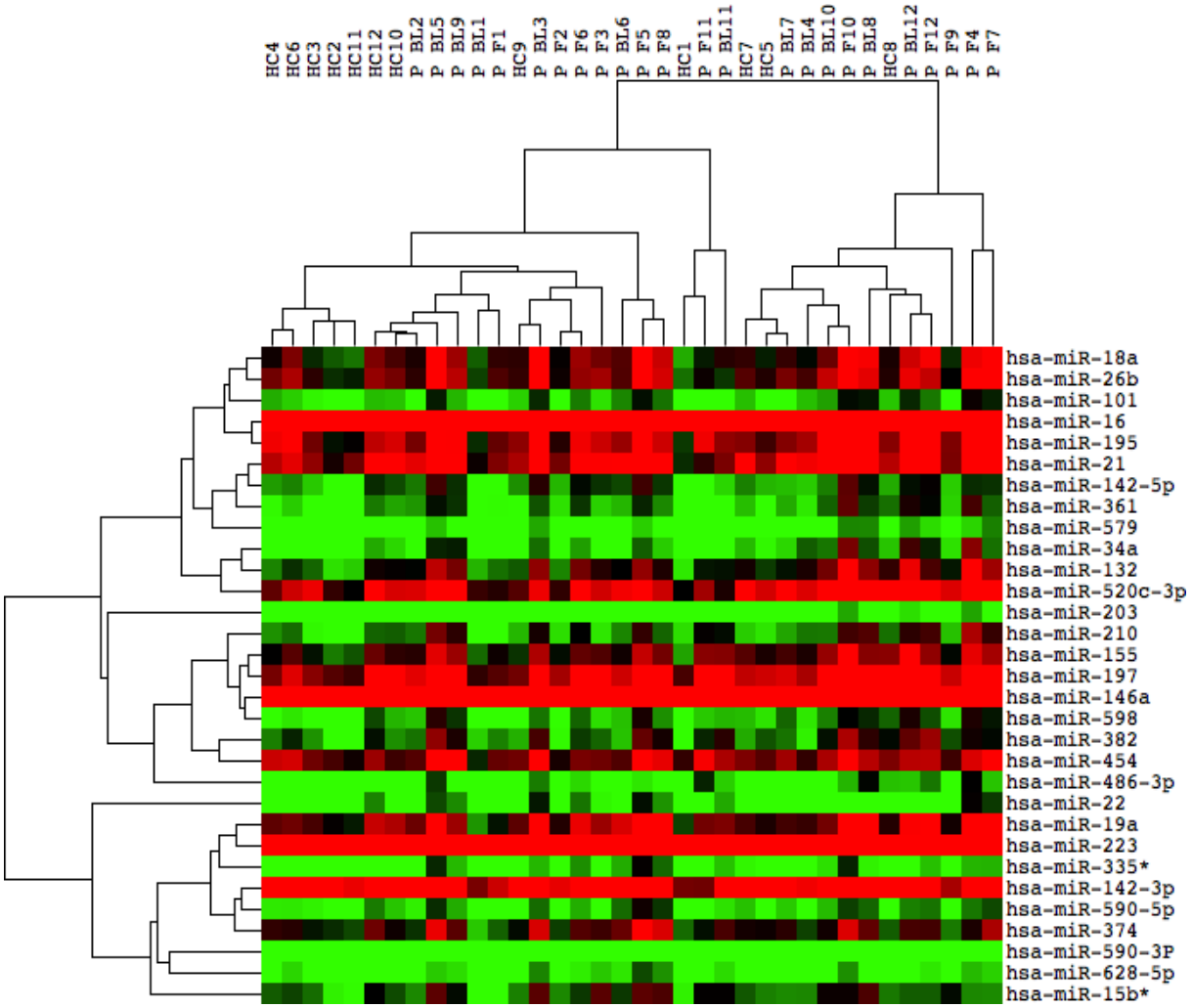


Figure 24 Heatmap for validation phase.

Healthy cohort (HC) and CCP progressors baseline (P BL) and CCP progressors follow up samples (P F). MiRNAs heatmaps were generated using hierarchical clustering (Gene Cluster 3.0 and Java TreeView) Green indicates low expression; Red indicates high expression levels.

Table 22 Within-patient changes in 31 miRNAs of interest for CCP+ patients in the validation phase.

		Validation phase						Pilot phase from Table18	
		CCP+ non-progressors			CCP+ progressors			CCP+ progressors	
	miR	Median (IQR) ddCt, mean (SD)	FC	N upregulated (/12)	Median (IQR) ddCt, mean (SD)	FC	N upregulated (/12)	Median FC	N upregulated (/12)
1	miR-16	-0.3 (-1.5, 0.3), -0.7 (1.4)	1.2	8	-0.9 (-2.1, 1.3), -0.7 (2.3)	1.9	7	1.3	10
2	miR-18a	-0.1 (-1.0, 0.1), -0.5 (1.3)	1.1	8	-0.1 (-1.8, 0.7), -0.4 (1.7)	1.1	6	3.1	10
3	miR-19a	-0.4 (-1.4, 0.2), -0.7 (1.3)	1.3	7	-0.2 (-1.9, 0.6), -0.5 (1.8)	1.2	7	1.5	9
4	miR-21	-0.5 (-1.3, 0.5), -0.5 (1.4)	1.4	8	0.0 (-1.4, 1.1), 0.0 (1.5)	1.0	6	1.6	9
5	**miR-22**	**-1.8 (-3.6, 1.1), -2.3 (1.7)**	**3.4**[†]	**12**	**-1.3 (-4.3, 1.1), -1.8 (3.6)**	**2.5**[†]	**8**	**4.3**	**12**
6	miR-26b	-0.3 (-1.1, 0.4), -0.3 (1.3)	1.3	7	-0.4 (-1.9, 0.4), -0.5 (1.6)	1.3	7	1.7	10
7	miR-34a	-0.4 (-1.7, 0.9), -0.4 (1.8)	1.3	7	0.3 (-1.9, 1.6), -0.1 (1.9)	-1.2	6	1.1	6
8	miR-101	-0.4 (-1.8, 0.1), -0.8 (1.3)	1.3	8	-0.1 (-1.8, 1.0), -0.3 (1.8)	1.1	7	2.1	11
9	miR-132	-0.5 (-1.2, 0.3), -0.4 (1.3)	1.4	8	0.3 (-1.7, 1.4), -0.2 (1.8)	-1.2	5	1.7	11
10	miR-142-3p	-0.1 (-1.3, 0.6), -0.2 (1.5)	1.1	7	-0.1 (-1.1, 1.2), 0.2 (1.4)	1.1	6	1.4	10
11	miR-142-5p	-0.6 (-1.2, 0.3), -0.4 (1.3)	1.5	7	-0.1 (-1.3, 0.7), -0.3 (1.3)	1.1	6	2.4	10
12	miR-146a	-0.5 (-1.3, 0.0), -0.5 (1.4)	1.4	9	0.1 (-1.0, 1.1), -0.2 (1.5)	-1.1	6	1.4	8
13	miR-155	-0.5 (-1.2, 0.4), -0.3 (1.2)	1.4	8	0.1 (-1.3, 0.9), -0.2 (1.4)	-1.1	6	1.2	8
14	miR-195	-0.2 (-1.4, 0.4), -0.6 (1.4)	1.2	7	-0.1 (-2.0, 1.4), -0.5 (2.2)	1.0	6	2.1	11
15	miR-197	-0.6 (-1.3, 0.2), -0.5 (1.2)	1.5	8	0.5 (-1.1, 0.9), -0.1 (1.5)	-1.5	5	1.5	7
16	miR-203	-1.6 (-2.8, 0.9), -0.9 (2.8)	3.1[†]	9	0.1 (-2.1, 1.8), -0.2 (2.6)	-1.1[†]	5	1.0	6
17	miR-210	-0.4 (-1.2, 0.1), -0.5 (1.2)	1.3	8	-0.2 (-1.9, 1.7), -0.1 (2.1)	1.2	7	3.4	10
18	miR-223	-0.7 (-1.6, 0.4), -0.5 (1.5)	1.6	8	-0.3 (-1.1, 0.4), -0.4 (1.2)	1.3	8	1.3	6/9

19	miR-361	-0.6 (-1.5, 0.5), -0.4 (1.3)	1.5	8	0.3 (-1.2, 0.9), -0.2 (1.6)	-1.2	5	2.9	11
20	miR-374	-0.5 (-0.7, 0.3), -0.3 (1.3)	1.4	7	-0.3 (-1.3, 0.6), -0.2 (1.5)	1.2	6	1.5	12
21	**miR-382**	**-1.3 (-1.4, 0.0), -0.7 (1.4)**	**2.4**	**9**	**-0.2 (-1.4, 1.1), -0.3 (1.6)**	**1.2**	**6**	**4.1**	**11**
22	miR-454	-0.1 (-0.9, 0.3), -0.2 (1.1)	1.0	7	-0.5 (-1.8, 1.1), -0.3 (1.6)	1.4	8	1.4	10
23	**miR-486-3p**	**0.0 (-1.6, 0.5), -0.5 (1.4)**	**1.0**	**6**	**-1.1 (-2.6, 1.3), -0.9 (2.8)**	**2.2**	**7**	**4.1**	**9**
24	miR-520c-3p	-0.8 (-1.4, 0.2), -0.4 (1.6)	1.7	9	0.3 (-1.1, 2.1), 0.4 (1.9)	-1.2	4	-1.3	5
25	miR-579	-1.7 (-1.9, 0.0), -0.8 (2.0)	3.2†	9	0.1 (-1.7, 1.6), -0.5 (2.5)	-1.1	5	2.2	11
26	miR-590-3P	-0.4 (-1.2, 1.9), 0.0 (2.1)	1.3†	6	-0.5 (-0.9, 1.3), -0.2 (1.8)	1.4†	7	1.9	8
27	miR-590-5p	-0.5 (-1.2, 0.5), -0.5 (1.4)	1.4	8	-0.4 (-1.3, 1.6), 0.1 (1.5)	1.3	7	1.4	10
28	miR-598	-0.5 (-1.0, 0.0), -0.5 (1.1)	1.4	9	0.2 (-1.2, 1.3), 0.0 (1.6)	-1.1	5	2.2	12
29	miR-628-5p	-0.3 (-0.8, 1.1), 0.1 (1.9)	1.2†	7	-1.0 (-1.9, 0.6), -0.8 (1.6)	2.0†	7	-1.3	5
30	miR-15b	-0.6 (-1.4, 0.0), -0.6 (1.3)	1.5	9	-0.1 (-1.2, 0.2), -0.3 (1.5)	1.1	7	6.1†	11/11
31	miR-335	-0.2 (-1.6, 0.7), -0.5 (1.6)	1.2†	8	-0.7 (-1.7, 0.6), -0.5 (1.7)	1.7	9	3.8†	12

MiRNA highlighted in bold met criteria of median FC≥4 and ≥75% consistent dysregulation in the pilot phase. The last two columns allow for comparison of results between the two phases. **†** *In these miRs, Ct was >32 for some patients at one or both time points and/or some values were undetermined and had to be imputed.*

5.4.3.5 Association of MiRNAs to clinical variables

Pooling results from both phases enabled associations of clinical variables for the 3 miRNAs of interest to be evaluated Table 23. MiR-382 and 486-3p were both associated (|rho|>0.3) with tender joint count and DAS28ESR. All three miRNAs were substantively associated with patient disease activity visual analogue score.

Table 23 Associations between clinical variables and dCt at baseline for key miRs

Baseline variable	miR-22	miR-382	miR-486-3P
Age	0.31	0.01	-0.10
EMS	-0.19	0.07	-0.10
TJC28	-0.23	-0.46	-0.39
Patient DA VAS	-0.33	-0.49	-0.65
Physician DA VAS	0.24	-0.31	-0.28
hsCRP	0.30	-0.03	0.01
ESR	0.05	-0.12	-0.19
DAS28ESR	-0.08	-0.40	-0.40

Values presented are Spearman's rho

EMS early morning stiffness, TJC28 tender joint count 28, DA VAS disease activity visual analogue score, CRP C-reactive protein, ESR erythrocyte sedimentation rate, DAS28ESR disease activity score using 28 joints and ESR

5.4.3.6 Predicting progression using baseline miRNA expression

As illustrated by Table 24, only miR-197 and miR-335#* performed better than chance (with 90% confidence) at predicting progression with AUC ROC 0.69 (90% CI 0.52, 0.85) and 0.71 (90% CI 0.52, 0.85) respectively. Considering the three miRNAs of interest; MiR-22 had a sensitivity and specificity of 63% and 100% respectively, with an AUC ROC curve 0.68 (90% CI 0.48, 0.82). MiR-22 did however result in a high Youden index, highlighting the importance of re-evaluating this in a larger sample size. MiR-382 and miR-486-3p performed less well reflecting in poorer sensitivity and specificity analyses and AUC ROC curve 0.57 (0.40, 0.75) and 0.55 (0.36, 0.72) respectively.

Table 24 Baseline miRNAs expression as predictors for progression.

miRNA	Progressors vs. non-progressors			
	AUC ROC (90% CI)	Sensitivity*	Specificity*	Youden J
miR-16	0.59 (0.40, 0.75)	33%	92%	0.25
miR-18a	0.55 (0.36, 0.72)	83%	42%	0.25
miR-19a	0.53 (0.36, 0.72)	83%	42%	0.25
miR-21	0.65 (0.44, 0.79)	50%	83%	0.33
miR-22	**0.68 (0.48, 0.82**)	**63%**	**100%**	**0.63**
miR-26b	0.53 (0.36, 0.72)	83%	50%	0.33
miR-34a	0.63 (0.44, 0.79)	75%	58%	0.33
miR-101	0.64 (0.44, 0.79)	83%	67%	0.50
miR-132	0.66 (0.48, 0.82)	50%	92%	0.42
miR-142-3p	0.53 (0.36, 0.72)	42%	75%	0.17
miR-142-5p	0.58 (0.40, 0.75)	67%	67%	0.33
miR-146a	0.65 (0.48, 0.82)	42%	92%	0.33
miR-155	0.63 (0.44, 0.79)	75%	58%	0.33
miR-195	0.63 (0.44, 0.79)	83%	50%	0.33
miR-197	0.69 (0.52, 0.85)	92%	58%	0.50
miR-203	0.60 (0.44, 0.79)	50%	83%	0.33
miR-210	0.57 (0.40, 0.75)	92%	42%	0.33
miR-223	0.60 (0.40, 0.75)	33%	92%	0.25
miR-361	0.67 (0.48, 0.82)	67%	75%	0.42
miR-374	0.51 (0.32, 0.68)	17%	92%	0.08
miR-382	**0.57 (0.40, 0.75)**	**75%**	**58%**	**0.33**
miR-454	0.53 (0.36, 0.72)	50%	67%	0.17
miR-486-3p	**0.55 (0.36, 0.72)**	**50%**	**75%**	**0.25**
miR-520c-3p	0.64 (0.44, 0.79)	58%	83%	0.42
miR-579	0.68 (0.48, 0.82)	83%	67%	0.25
miR-590-3P	0.54 (0.36, 0.72)	75%	50%	0.25
miR-590-5p	0.58 (0.40, 0.75)	83%	42%	0.25
miR-598	0.62 (0.44, 0.79)	33%	92%	0.25
miR-628-5p	0.55 (0.36, 0.72)	17%	100%	0.17
miR-15b#*	0.58 (0.40, 0.75)	83%	50%	0.33
miR-335#*	0.71 (0.52, 0.85)	50%	92%	0.42

*At cut-off that maximised Youden J (sensitivity +specificity-1), prioritising specificity if tied.

5.5 Discussion

As identified in the literature review (section 2.2.2) and subsequent chapters, several characteristics have been considered as biomarkers for within at-risk cohorts. However, this is the first miRNA study to be performed in individuals with autoimmunity at risk to progression to RA. Previously, research groups have focussed on identifying specific miRNAs highly and/or uniquely expressed in (established) RA (Table 17). More recently, there have been studies evaluating the change in expression with therapy [251] as well as whether miRNAs can predict response to treatment [366]. Given the heterogeneity of RA, it is not surprising that these studies have included cohorts that have often been poorly characterised. This at-risk study has aimed to improve the characterisation of the cohorts by limiting the inclusion criteria to those autoantibody positive individuals with no synovitis on ultrasound scan (as ultrasound synovitis can be considered immunopathologically closer to RA disease). This enabled the recruitment of a clearly defined cohort in whom miRNAs identified were to be of importance in disease initiation or, conversely, may offer a protective function. Furthermore, this study has enabled a unique opportunity to observe dynamics of miRNA expression associated with progression or non-progression to RA with the use of matched samples. The variability in laboratory and analytic techniques employed in miRNA studies makes reliability and comparison between studies challenging. Steps have been taken to ensure a robust analytical approach has been adopted through working closely with a biomedical statistician.

Global profiling of over 700 miRNAs in the pilot phase enabled a vast number of potentially influential miRNAs to be considered. The expression of miRNAs across the continuum from health to RA or VERA identified several highly expressed miRNAs. The hierarchical clustering analysis also suggested that the studied cohorts (HC, CCP+ and VERA) could be broadly distinguished on their miRNA expression. As might be expected, the number of miRNAs that are dysregulated is greatest comparing health to VERA (12 upregulated, 1 downregulated) as opposed to health and CCP+ (4 upregulated, 4 downregulated). Similarities between the two states compared to health were seen with miR-142-3p, -520c-3p, -590-3p and -628-5p which were upregulated and miR-197

which was downregulated in both CCP+ and VERA cohorts. However, the within patient change from CCP+ to VERA in matched samples for these miRNAs did not meet criteria. This may be a consequence of significant biological change already present in the CCP+ state. The results suggest that transition from health to a state of active inflammation (VERA) requires additional miRNAs to be activated compared to states of autoimmunity without inflammation (CCP+ arthralgia). Although disease onset and inflammation cannot be solely accounted for by the change in miRNAs reported, their association with inflammation requires further investigation in larger cohorts.

MiRNAs role in disease initiation was explored through the matched sample component of the pilot phase. The three miRNA (miR-22,-382, 486-3p) that met the study criteria were significantly upregulated from CCP+ to VERA and offer insights into a mechanistic role in RA initiation. However, due to logistical difficulties having two time point samples to measure change in expression, their clinical utility as biomarkers may be limited.

The validation phase enabled the 31 miRNAs of interest to be evaluated in a new cohort of matched samples. Establishing whether the changes identified in the progressors were unique to transition from CCP+ to VERA was achieved through observations in an additional cohort of CCP+ non-progressors. This cohort also allowed for insights into how miRNAs may be used as biomarkers. It is acknowledged that the non-progressors findings have not been validated. Comparisons across the cohorts resulted in 4 miRNAs being validated from the pilot phase. These were within the VERA to health comparisons and include miR-22 (FD 9.8). Similarly, in the comparison of CCP+ (progressors) and health, miR-22 was also upregulated (FD 11.3). Notably, upregulation in the non-progressor CCP+ group was not demonstrated. This provides further evidence to support miR-22 role in disease initiation and inflammation. In order to ascertain whether this miRNA could have a biomarker role in predicting progression, sensitivity /specificity and area under the curve ROC analysis was performed. Despite a specificity of 100% (sensitivity 63%) and a moderately good AUC ROC score, the 90% confidence interval crossed 0.5 and therefore suggests it is no better than chance at predicting progression. This may be due to the small sample size and certainly requires further examination in a larger cohort in light of the good Youden Index score. Of the remaining

signature miRNAs from the pilot phase, miR-382 showed similar expression while miR-486-3p had higher expression in the non-progressors. These findings are difficult to interpret given the small sample size and warrant further review in a larger cohort.

In the validation matched analysis, expression levels of the signature miRNAs did demonstrate consistent upregulation in the progressors. However, these did not meet the study's predefined criteria. Initially adopting such stringent criteria allowed highly dysregulated miRNAs to be identified and assisted in filtering out intermediary levels. MiR-486-3p and MiR-22 had a FC of ≥2 in the validation phase – it should be noted that this is a value that is consider to be of biological significance in the field and therefore should not be discounted.

The role of this new cohort in the validation phase was to ascertain the change in expression levels of the miRNA of interest. It was hypothesised that the miRNA with the greatest levels of change in the progressor cohort would demonstrate stable expression in the non-progressor cohort. In fact, the results indicated that two of the three key miRNAs were upregulated to some degree in the non-progressors. One can speculate that these miRNAs are therefore upregulated throughout the cohort over time and are a feature of autoimmunity (CCP positivity) and at-risk states rather than being associated with progression to inflammatory disease. Furthermore, recognition of the dynamic and evolving nature of this cohort of individuals is important. Although no progression had occurred during observation period, individuals remain on a continuum and there is the possibility that they may evolve to inflammatory disease in the future and hence some upregulation is plausible.

Although this analysis has not had sufficient sample size to allow for inclusion of miRNAs within a predictive model, association with clinical variables has been considered. All of the 3 miRNA identified in the pilot phase were associated with the patient reported disease visual analogue score. MiR-382 and -486-3p were also associated with tender joint and disease activity scores (DAS28ESR). Interpreting these finding remains difficult.

However, the fact associations exist with other clinical variables that assist in the diagnosis of inflammatory arthritis is encouraging.

The miRNAs previously reported to be associated with inflammation (miR-146a and -155) were found to be upregulated in our pilot study, although not sufficient enough to be part of the predictive signature. There was negligible upregulation in the validation phase. This is perhaps unsurprising as it is the first time a matched pre-RA to RA population has been studied, and we would anticipate that different miRNAs might be involved in disease initiation compared to previously studied established disease cohorts. Furthermore, techniques to measure miRNAs in sera have only been established relatively recently. Previous studies have predominantly focused on synovial tissue/fluid and whole blood expression rather than sera. Therefore, miRNA in the sera may be unlikely to be the same as those found within the joint.

It is apparent from reviewing the literature that analysis of pre and post progression to disease data are limited. This is undoubtedly due to the lack of availability of such unique samples. Studies of other 'pre-disease' states have considered the utility of miRNAs in differentiating at-risk individuals e.g. Barrett's oesophagus prior to development of oesophageal malignancy [253]. These have been cross sectional and no study investigating the change in miRNA expression in an individual pre- and post-disease onset. Within the field of autoimmunity, a recent cross-sectional project evaluated the expression of MiRNAs associated with development of lymphoma in primary Sjogrens syndrome (PSS) [370]. In this study, 12 PSS patients with lymphoma, 12 PSS patients without lymphoma, and 12 healthy controls were considered. MiRNA array profiling revealed a clear clustering of the 3 subject groups with 44 dysregulated miRNAs. The expression levels of these 3 miRNAs (with the highest FC) enabled sufficient differentiation of PSS patients with lymphoma from those without. Certainly, for identifying biomarkers this is the optimum methods to adopt. In comparison, the use of matched samples would provide insights into potential mechanistic processes involved in disease initiation.

The possible functional role of the three miRNAs identified has been preliminary assessed using the bioinformatics platform MetaCore™ which considers pathway prediction for miRNAs (LO conducted analysis). The expanded networks generated for miRNAs of interest represent predicted targets. Canonical interaction between the transcription factor p53 and miRNA-22 is highlighted; p53 plays a central role in a number of cellular functions, and is overexpressed in RA synovial tissue, and also activates miR-22 by binding to its promoter region. Predicted network shows that MiR-486-3p has an inhibitory effect on the bone morphogenetic protein 1 (BMP-1), indicative of miRNA function. MiR-382 negatively regulates the phosphatase and tensin homolog PTEN, which is upstream of the AKT/mTOR signalling pathway. This analysis assists development of potential functional work focusing on these miRNAs.

MiR-22 has been highlighted as a potential biomarker of disease, whilst its change in matched samples suggests a possible mechanistic function. MiR-22 was originally reported as a tumour supressing miRNA [371, 372]. Recently, studies have reported its role in oncogenic disease development of several cell lines including those associated with lung, colorectal and hepatocellular carcinoma [373-375]. Its pathway response with the transcription factor p53 may account for these functions. The tumour suppressor protein p53, plays a central function in cell cycle regulation, DNA repair, and apoptosis and has been shown to engage miRNAs for tumour suppression including miR-22 [376]. Somatic mutations in p53 and its overexpression has been observed in RA synovial tissue [377, 378]. There are no studies reporting on serum miR-22 levels in RA, however one group has identified an interaction of p53 and miR-22 with regulation of Cyr61. In synovial samples the secretion of the extracellular matrix protein Cyr61 by FLS has been shown to i) contribute to the hyperplasia of synovial lining and ii) perpetuate the IL-6/Th17 inflammatory cycle [379, 380]. Subsequently, a novel mechanism has been identified in which Cyr61 production is regulated by p53 via its activation of miR-22 [381]. MiR-22 was shown to target and inhibit Cyr61 expression, resulting in the negative correlation with Cyr61. The authors demonstrate that in RA, mutant forms of p53 are unable to activate miR-22 transcription. The reduction in miR-22 leads to uninhibited expression the Cyr61 contributing to proliferation of FLS and inflammatory Th17 pathways. Within this study low levels of miR-22 are reported at the

synovium. As previously alluded to, it is difficult to correlate finding from synovial level studies in established RA to that of serum studies in at-risk individuals. However, the potential of miR-22 in the pathogenesis of RA supports further validation within at-risk cohorts.

5.6 Limitations

When working with a unique study cohort one of the most challenging aspects is devising a project which provides sufficient numbers for statistical or meaningful analysis. The modest numbers and subsequent limitations is acknowledged. This project has involved a well-phenotyped group of individuals with positive autoantibodies, no clinical synovitis, and in addition no ultrasound evidence of synovitis (as deemed by absence of PD). The numbers included in the pilot phase reflect common practice for exploratory phases of studies. However, a greater number of samples would have ideally been incorporated in the validation study. At the time of random selection of samples the number within the cohort with no Power Doppler on baseline imaging was 91/136 (66.9%). This included both individuals that progressed and those that did not. Within the progressor group (n=57 of the 136 (42%)) only 29 of these individuals had no PD at baseline. The availability of samples for eligible individuals was a further limitation as specific time points were occasionally not available. This limited the number of individuals eligible for random selection and entry into the validation phase.

In the validation phase, imputation of undetermined Ct values was performed by the departmental statistician. The authors of this statistical software package used reported how undetermined values of >40 were actually miRNAs that failed to amplify rather than lack of presence. The package uses the observed values in the dataset to determine the likely undetermined values. This allowed a mean value of all the Cts within the triplicate to be determined. This may be considered a possible solution to handing undetermined values.

This study followed the accepted format for miRNA studies with a pilot phase evaluation of a large number of miRNA via global profiling followed by a focused validation phase. Global profiling generates a vast amount of data and is an expensive approach not feasible for larger clinical cohorts. Therefore, for validation, custom cards plates are frequently used in which a selection of the miRNAs of interest are evaluated. Use of custom cards and difference in normalisation techniques may account for the differences found in expression of miRNAs in the validation phase. In the pilot phase, normalisation was achieved by calculating the average value across total expressed miRNAs for each sample (Normfinder software). Whereas, with custom cards in validation phase, the manufacturers' recommendation to use the mean of the 3 endogenous control replicates (RNU6) was applied. It is acknowledged that in this and similar translation projects the analytical approaches can be very disparate. In order to ascertain the degree of correlation between these normalisation methods, future studies are required.

Use of global profiling and then custom cards is common place in PCR/miRNA studies. The manufactures of the custom cards and the global profiling plates are the same and hence deemed transferable. However, in this study, no specific quality control experiment was performed to ensure uniformity in the two types of plates used. Ideally, additional experiments should have been conducted allow for several pilot phase samples analysed on the global profiling cards to be run on the custom cards and vice versa for the validation phase. This would produce ddCts for each miRNA of interest generated using both methods. This would allow for direct comparisons and quality control. It would be expected that a moderate degree of correlation would be achieved between the two methods. Retrospective analysis may not be beneficial given that experimental conditions may fail to replicate those at the time of the initial study e.g. age of reagents, quality of samples.

Alternatively, the limitations listed above could be eliminated by applying global profiling cards in both phases, thus enabling the same normalisation method and cards to be used. However, this would have considerable financial implications and it would

not be realistic to anticipate clinical utility of such a test to be made transferable at such costs.

Despite efforts to ensure the cohorts between phases were similar clinically, the variation in methods to analyse miRNA expression in the two phases may account for the difference in expression.

5.7 Key points

- Individuals can be broadly distinguished by their serum miRNA expression from health to at-risk (CCP+) and on to early arthritis (VERA), although expression remains varied.
- The use of matched samples prior to and at point of inflammatory arthritis enabled insights into the mechanistic role of specific miRNA. The 3 miRNA identified in the pilot phase failed to meet the study criteria in the validation phase. However, miR-486-3p and miR-22 did have a FC >2, a potentially biologically significant result.
- Within a cohort of at-risk individuals (CCP+ progressors and non-progressors), MiR-22 has potential as a biomarker to differentiate those at risk to arthritis development.

5.8 Conclusions

The findings and limitations identified by this work demonstrates the challenges common place in translational research. It is acknowledged that despite attempts to accurately phenotype individuals, there is likely to still be heterogeneity within cohorts. It is also apparent that the inconsistencies in experimental and analytical methods within this area may impact the validity of these and other published results. At this time, an exact miRNA signature cannot be easily defined. Instead there are several miRNAs which are expressed during the CCP+ state and on through to progression to VERA. This is the first study which has adopted a comprehensive miRNA array method

to consider miRNAs in an at-risk cohort. The use of matched samples has allowed for 3 miRNAs to be identified that appear to be associated with autoimmunity (at risk to RA) and the progression to RA inflammation. In particular, miR-22 potentially has utility in identifying which individuals within an at-risk cohort may progress to inflammatory arthritis.

Further clarification and validation of these findings will be possible through analysis of MiRNA expression in larger cohorts alongside additional functional work of the candidate miRNAs. A further area of interest would be to widen the inclusion criteria to include those individuals with autoimmunity and ultrasound evidence of synovitis but not clinical detectable synovitis. This may not have a biomarker role but perhaps offer insights into the changes occurring along the IA continuum. It would be expected that individuals with ultrasound detected synovitis would have a similar miRNA profile as those with VERA and provide further evidence that these are individuals that warrant therapy.

6. T-cell subsets in individuals with systemic autoimmunity and arthralgia: an immunological biomarker.

This chapter describes the immunological phenotyping of individuals within the at-risk cohort using T-cell subset. Use of a prospective study design allowed statistical modelling to be performed to quantify and qualify whether T-cell subsets predicted progression to IA. Work featured in this chapter has been peer reviewed and published [382].

6.1 Introduction

The exact pathogenesis of RA remain uncertain, although immune cells and their interactions are thought to be essential [383]. The affiliation of disease with autoantibodies and the mechanisms for major histocompatibility complex linkage support the T-cell driven concept. The demonstration of multiple immune cells within the RA synovium, particularly high levels of both CD4+ and CD8+ T-cells, further endorse the potential pathogenic role of T-lymphocytes [384-386]. More recently, the identification of T-cell related genes and disease response to T-cell modulation therapy has strengthened this view point [383, 387-389]. As research in the field has evolved, there is mounting evidence to suggest that T-cells are not the only mechanism of disease initiation and perturbation [390].

T-cell subset quantification is one method of gaining insight into the immune status of patients with RA [391]. Of the subsets, regulatory T-cells (Treg), defined as CD4+CD25+FOXP3+ have been the focus of many studies [247, 392-394]. Their capability of suppressing T-cell activation, proliferation and effector function facilitates the maintenance of self-tolerance and regulation of autoimmune responses. In synovial fluid, studies have demonstrated increased Treg cell numbers, suggesting a migration and expansion of Tregs to the synovium. In peripheral blood studies the findings have been less consistent, although differences in phenotype of RA individual evaluated (early disease vs established and untreated vs DMARD), may account for some of the

inconsistencies [247, 394-396]. Furthermore, as experience within the field has grown, the cell markers used to define cell populations have evolved. A highly purified population of Treg cells are can be identified using cell markers CD4+, $CD25^{high}$, Foxp3+, $CD127^{low}$ [397, 398].

Inflammation has a direct effect on T-cell differentiation in RA. Work from Leeds and other centres has focussed on the significance on CD4+ T-cells including naïve T-cells [399]. It is postulated that inflammation acting on the thymus causes a reduction in circulating naïve cells [399, 400]. Initial findings demonstrated a reduction of both naïve and regulatory CD4+ T-cell frequency in the peripheral blood of individuals with active RA [247, 399]. Evaluation of T-cell subsets in patients along the IA continuum has demonstrated naïve CD4+ T-cells to be an important T-cell biomarker for treatment response [248-250]. Sustained remission was associated with low levels of immunological dysregulation (as defined by higher proportions of naïve and Treg cells), suggesting that once immunological equilibration is achieved successful withdrawal of therapy is possible [249]. Specific levels of T-cells subsets have been shown to predict i) methotrexate-induced remission [250], ii) relapse following DMARD-induced remission [248], and iii) discontinuation of TNF-blockers in RA individuals without subsequent flare [249].

In addition to Treg and naïve cells, a novel T-cell subset expressing both naïve and memory differentiation markers was noted in the peripheral blood of RA individuals. Retention of cell markers CD45RA and CD45RB suggested that these were immature T-cells. These so-called, inflammation-related cells (IRC), occur following differentiation of naive T-cells into other subsets influenced by pro-inflammatory cytokines such as IL-6 and TNF [399]. Other groups have reported a similar subset and demonstrated through functional work that they can express chemokines receptors such as CXCR4 [401-403]. In active RA, IRC frequency in the peripheral blood correlates directly with systemic levels of inflammation as measured by CRP [399]. In vitro studies have demonstrated that IRCs isolated from synovial samples express chemokine receptors (that direct them towards inflamed tissue when disease is active [248]. Unexpectedly, their presence has also been observed in patients with RA in remission [248]. In these

patients, however, the IRCs did not proliferate or act as responsively as in active disease. In this population, their presence was associated with flare of disease within 18 months. It is proposed that in the remission setting, detectable IRCs in the peripheral blood are a product of T-cells released from the synovium which have a resistance to apoptosis. They remain in the circulation and are primed for reactivation with the next inflammatory trigger or event. Their measurement and prognostication value has not been assessed in individuals at risk of IA/RA.

This programme of work considers the potential of T-cell subsets (consisting of naïve, Treg and IRC) as a biomarker to predict progression to IA in at-risk individuals.

6.2 Aims and overview

The aim of this study was to report on the extent of T-cell subset dysregulation in CCP+ individuals with non-specific musculoskeletal symptoms (from a single baseline blood sample) compared to healthy individuals. Furthermore, the study looked to establish whether dysregulation was associated with progression to IA. Given that a clinical model for progression to IA has already been reported (Section 2.3 and [217]), it was important to determine the confounding effect of clinical parameters on any T-cell model that was generated. It was hypothesized that in CCP+ individuals with non-specific symptoms, those with the greatest T-cell subset dysregulation (as determined using naïve CD4+T-cells (naïve), inflammation related cells (IRC) and regulatory T-cells (Treg) quantification) would have a greater propensity for progression to inflammatory arthritis. Immune dysregulation is one of the primary events in disease onset. It was therefore hypothesized that an early predictor of disease progression could be dysregulated T-cell subsets.

6.3 Patients and Methods

6.3.1 Patients

Individuals who were CCP+ and had non-specific musculoskeletal symptoms were recruited for the study as detailed in the methods chapter (Section 3.4 Study Population). One hundred and three participants were recruited. Inclusion required both a valid sample for T-cell subset quantification at baseline and prospective follow-up for over 6 months duration. One hundred and six individuals (known to be anti-CCP negative) formed a healthy control group for comparison.

6.3.2 Clinical Assessments:

All participants provided baseline demographic details, patient questionnaires, and clinical history of symptoms and had a clinical examination by a rheumatologist (LH & CR) which included a joint count. Individuals then followed the study schedule as outlined in Section 3.5 Study Protocol. Individuals attended 3 monthly visits for the first year and as clinically indicated thereafter for up to 6.5 years. Participants were able to attend in between visits if they developed any new symptoms. A number of blood tests were undertaken at baseline including inflammatory markers and HLA status. HLA-DRB1 shared epitope status (low-resolution) was considered positive in the presence of one or two copies of the following alleles: HLA- DRB1*01, DRB1*04 and DRB1*10 in the HLA-DRB1 locus [404, 405]. All the clinical parameters that had been used in the initial clinical prediction model were performed. However, baseline ultrasound data was not available for all the participants.

6.3.3 Laboratory Methods:

6.3.3.1 T-cell subset analysis:

Peripheral blood (up to 6mls) was collected into standard vacuettes containing ethylene diamine tetracetic acid (EDTA) solution and processed for red cell lysis within 2 hours of collection. Following red cell lysis, six-colour flow cytometry was performed on peripheral blood mononuclear cells using several cell markers (Box 5 & Figure 25). The antibodies used for each panel are listed below; identification of naïve and IRC CD4+T-

cell subsets required CD45RB-FITC (clone MEM-55, Serotec, Oxford, UK), CD45RA-PE (clone F8-11-13, Serotec), and CD62L-APC (clone 145/15 Coulter, High Wycombe, UK). Tregs quantified by cell surface marking for CD4-Pacific blue (clone RPA-T4, BD, Oxford, UK), CD25-APC (clone 2A3, BD) and CD127-PE (R34.34, Beckman coulter), followed by intracellular staining for FOXP3-FITC (clone PCH101 eBioscience, San Diego, CA) using the anti-human Foxp3 staining kit (Insight Biotechnology, Wembley, UK).

The flow cytometry analysis was performed on a LSRII cytometer (BD), using BD Biosciences FACSDIVA software. Subset frequencies were reported as percentage of gated CD3+/CD4+ T-cells.

Flow cytometry gating was performed by a senior colleague (FP).

Box 5 T-cell Subset flow cytometry markers

T-cell subset	**Cell markers**
Naïve CD4+T-cells	CD4+, CD45RB^{high}, CD45RA+, CD62L+
T-regulatory cells	CD4+, CD25^{high}, Foxp3+, CD127^{low}
Inflammation Related Cells	CD4+, CD45RB^{high}, CD45RA+, CD62L–

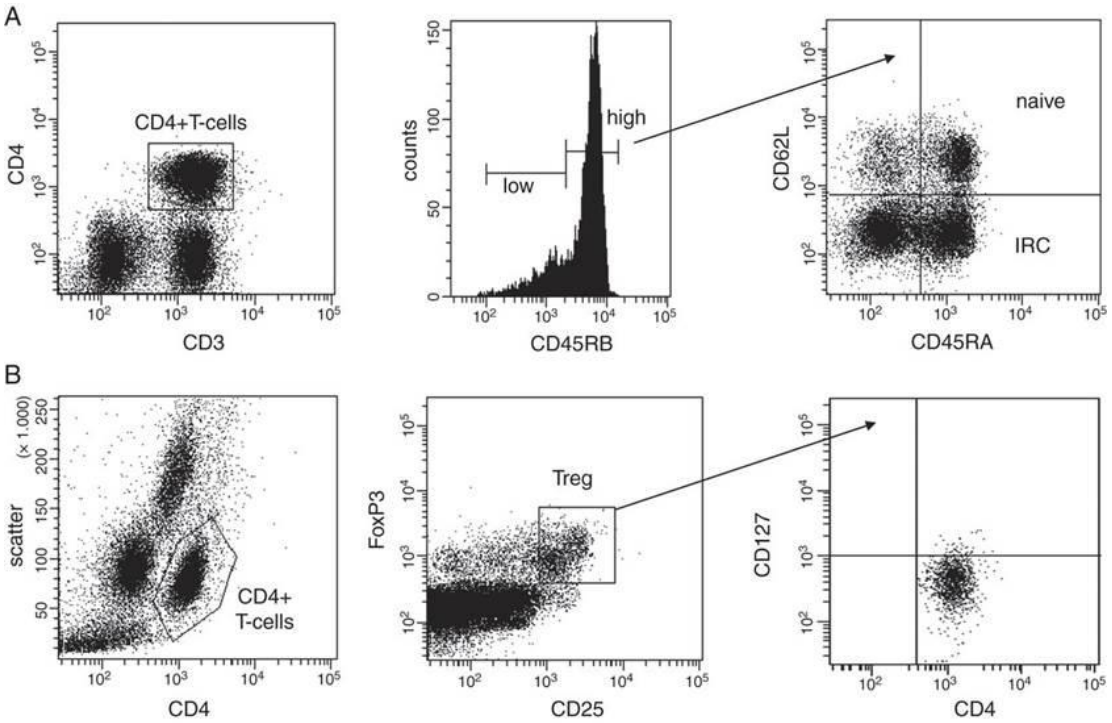


Figure 25 Flow cytometry gating strategies

Reprinted by permission from BMJ Publishing Group Ltd [250].

The top panels represent flow cytometry analysis of naïve cell & IRC frequency. For naïve cell and IRC subsets, CD4+ T-cells were gated using double positivity for CD3/CD4. High expression of CD45RB was used to gate CD4+/CD45RB^{high} cells and the gate was applied to a dual plot of CD45RA and CD62L. The bottom panels represent flow cytometry analysis for Tregs. CD4+ T-cells were gated using scatter properties and CD4 levels. Tregs were quantified using high expression of CD25 and Foxp3. The Treg phenotype was confirmed by documenting a low level of CD127 expression.

6.3.4 Statistical analysis:

Reference limits for the T-cell subsets

Reference limits for each T-cell subset were calculated (lower 95% limit of normal for naïve cells and Tregs, upper 95% limit of normal for IRC) using data from 106 healthy controls. This involved multiple linear regression to assess whether T-cell subset frequencies varied with age or sex. Where associations with age were found, one-sided 95% prediction intervals for the association were obtained by calculating two-sided 90%

intervals and discarding the upper or lower interval accordingly; a 90% confidence interval around the prediction interval was calculated. Otherwise, the 5% or 95% centile and 90% confidence interval were calculated. T-cell subset frequencies found to be skewed were log transformed prior to analysis. Back-transforming to the original units yielded asymmetric confidence intervals.

Description of T-cell subsets compared to health:

The development of reference limits allowed the dichotomisation of T-cell subset values as normal or abnormal. One-sample binomial tests were used to assess whether the proportion with abnormal value differed from the expected 5%. Analysis was restricted to patients with complete T-cell subset data. Pearson's Chi-square test was used to identify associations between shared epitope and T-cell subset abnormalities.

Unadjusted associations between T-cell subset frequencies and progression to IA:

Non-parametric area under the receiver operator curve (ROC) was calculated for each subset. Additionally, sensitivity and specificity of each subset for predicting progression to IA at any time during follow-up were calculated, with 95% confidence intervals estimated by the Wilson method.

Adjusted associations between T-cell subset frequencies and progression to IA:

Binary logistic regression models of the occurrence of progression to IA and Cox proportional hazards models of time to progression were constructed to adjust for the following variables, 1) age; 2) SE status (negative/positive); 3) smoking (never/ever); 4) CCP titre. Clinical variables from a previously published model were also considered [217]. These included, physician assessed small joint tenderness (absent/present) and duration of early morning stiffness (<30 minutes/≥30 minutes). Models were produced sequentially to investigate the effects of adding in covariates.

An adjusted model containing only the T-cell subsets and age was specified (model 1). This was compared to the published clinical model (model 2) and subsequently a

combined model consisting of both clinical and T-cell subsets (model 3). Analysis was first performed on the subset of patients with complete data to test model performance. Akaike Information Criteria (AIC) values were used to compare different models (lower AIC value indicates a superior model). To account for any missing data, analysis was repeated following multiple imputation with chained equations in 20 complete datasets. The general results were combined according to Rubin's rules. Predictive mean matching was used as the imputation model for all continuous variables; for binary variable, logistic regression was used.

The predicted probability of progression obtained from the final logistic regression model (model-3) was calculated for patients with full data. Patients were then categorised as being at low (<20%), moderate (20-80%) or high risk (>80%). Kaplan-Meier plots and associated log-rank tests were then produced for time to progression, using these risk groups.

Power calculation

Previous studies from the literature informed the numbers required for development of clinical reference ranges/limits [406]. For the logistic regression modelling of progression to IA the same statistical approach was taken as the previously published clinical model [217]. Guidance recommends that there are at least 10 cases in the smallest outcome category ('events') per variable (EPV) [407]. Subsequently, the final analysis was performed once there was sufficient follow-up and hence 'events', in this case progression to IA. It was proposed that T-cell abnormality would form a single variable alongside the previously established variables – early morning stiffness, joint tenderness and >3 x upper limit of normal antibody titre. Therefore at least 40 events were required.

6.4 Results

6.4.1 Patient Characteristics

Forty eight of the 103 (46.6%) patients developed synovitis during follow-up, with the majority of individuals, 30/48 (62.5%) progressing within 12 months. Baseline clinical and demographic data are presented in Table 25.

Table 25 Clinical characteristics of CCP+ individuals.

Characteristic		Result
Progressed (ever): % (n)		46.6% (48)
Duration of follow-up (months) median (range)		18.4 (0.1 to 79.6)
Age (years) mean (SD; range)		52.6 (11.7; 27 to 79)
Female: % (n)		71.8% (74)
SE positive* : %(n)		73.5% (72)
High positive CCP+ and/or RF ‡: % (n)		85.4% (88)
Smoker: % (n)		
	Non-	30.1 (31)
	Ex-	41.7 (43)
	Current	28.2 (29)
EMS≥30 mins: % (n)		34.0% (35)
Small joint symptoms: % (n)		43.7% (45)

*CCP cyclic citrullinated peptide; EMS early morning stiffness; IRC inflammation-related cells; RF rheumatoid factor; SE shared epitope. *available in 98/103 patients. ‡determined as >3 X upper limit of normal*

6.4.2 Reference Limit

Samples from 106 healthy controls enabled the development of reference limits for each T-cell subset. The following section details the development of the reference limits. Descriptively, naïve cell frequency was lower in older HC [399] but did not differ by gender (Figure 26 and Table 26). IRC were not related to demographic parameters (Figure 27). A clear positive association was found between Treg frequency and age (Figure 28, Table 27). There was no difference by gender.

Naïve cells

Naïve cell frequency was available for 106 controls; mean (SD) age 43.54 (12.52), range 19 to 69. There was no significant evidence that naïve cell frequency differed between males and females [age-adjusted difference (95% CI) -1.44% (-4.77%, 1.89%); p=0.392], or that its association with age differed by sex [difference in slope -0.02% (-0.24%, 0.27%) per year; p=0.910]. However, there was a highly significant tendency for naïve cell frequency to be lower in older people [slope -0.54% (-0.67%, -0.42%) per year; p<0.001]. The reference limit was therefore adjusted for age but was not stratified by sex.

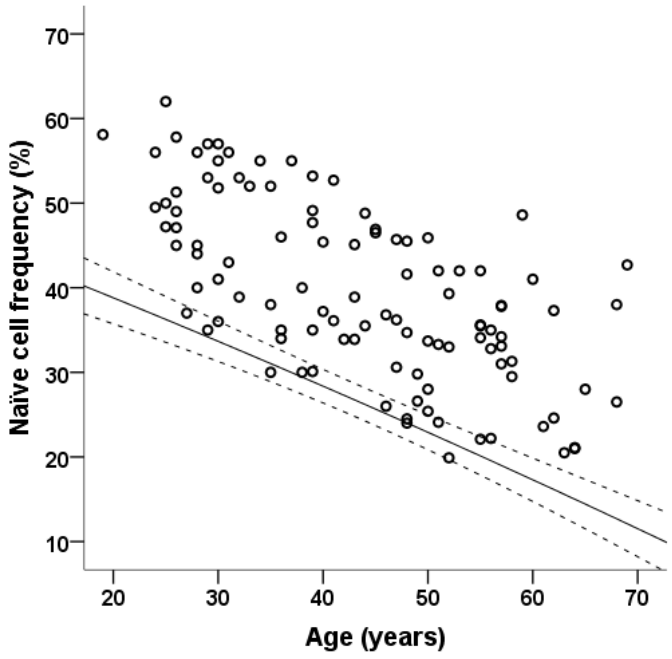


Figure 26 Scatter plot of naïve cell frequency (% of CD4+ T-cell) and age.

Solid line = Lower limit of normal, Dashed line = 90% CI

To calculate a one-sided 95% prediction interval, a two-sided 90% interval was calculated and the upper limit discarded. The naïve lower limit of normal (LLN) and corresponding 90% confidence interval around it were calculated as LLN.

Table 26 Lower limit of normal naïve cell frequency for ages in the range 25 to 65 years

Age	Lower limit of normal (90% CI)
25	36 (34, 39)

30	34 (31, 36)
35	31 (29, 33)
40	28 (26, 30)
45	26 (24, 28)
50	23 (21, 25)
55	20 (18, 22)
60	17 (15, 20)
65	14 (12, 17)

Outside this range of ages, it is possible that the association between age and naïve cell frequency may not be linear; therefore, more data is required for controls aged under 25 or over 65. The lower limit for these age groups are not provided.

Inflammation-related cells (IRC)

IRC frequency was available for 101 controls; mean (SD) age 43.50 (12.69), range 19 to 69. Data was log transformed prior to analysis. There was no evidence that IRC differed between males and females [geometric mean ratio 0.98 (0.68, 1.41); p=0.900] or varied with age [change -0.41% (-1.73%, 0.93%) per year; p=0.544]. The calculated 95% centile and its 90% confidence interval (back-transformed to original units) were found to be 3.70 (3.30, 7.00). This corresponded to the upper limit of normal for IRC.

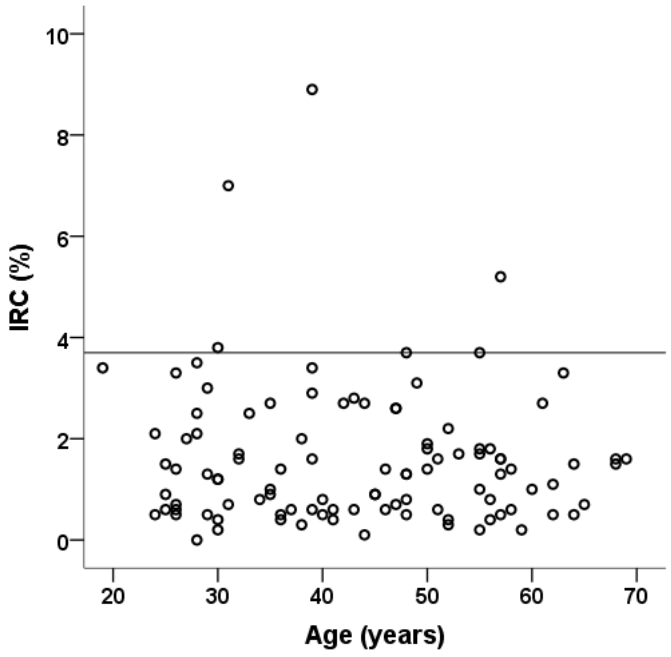


Figure 27 Scatter plot of IRC cell frequency (% of CD4+ T-cell) and age.
Solid line represents the upper limit of normal.

T-regulatory cells (Tregs)

Treg cell frequency was available for 98 controls; mean (SD) age 44.09 (12.30), range 19 to 69. There was no evidence that T-regulatory cell frequency differed between males and females [age-adjusted geometric mean ratio 1.03 (0.82, 1.22); p=0.976], or that its association with age differed by sex [ratio of differences in slope 1.00 (0.99, 1.02); p=0.677]. However, there was a statistically significant tendency for Treg cell frequency to be higher in older individuals [by 1.22% (0.50%, 1.94%) per year; p=0.001]. The reference range was therefore adjusted for age but was not stratified by sex.

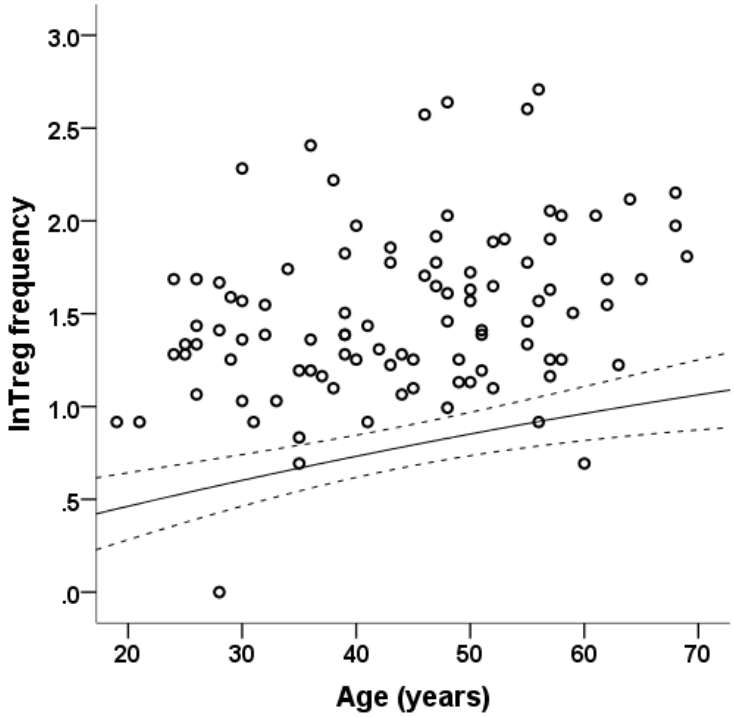


Figure 28 Scatter plot of ln-transformed Treg frequency (% of CD4+ T-cell) and age.

Solid line = Lower limit of normal, Dashed line = 90% CI

Table 27 Lower limit of normal Treg cell frequency for ages in the range 25 to 65 years

Age	Lower limit of normal (90% CI)
25	1.71 (1.46, 2.00)
30	1.83 (1.59, 2.10)
35	1.95 (1.72, 2.21)
40	2.08 (1.86, 2.33)
45	2.21 (1.98, 2.47)
50	2.34 (2.09, 2.63)
55	2.48 (2.18, 2.82)
60	2.62 (2.26, 3.03)
65	2.76 (2.33, 3.25)

This data allowed dichotomisation of values observed in CCP + individuals as within or below the normal range for naïve T-cell and Treg (defined by the lower limit of normal (LLN)) and within or above normal range for IRC (defined by the upper limit of normal (ULN)).

6.4.3 Unadjusted T-cell analysis

There were no abnormal lymphocyte counts in this cohort of individuals. The CD4+ T-cell numbers did not differ between progressors and non progressors.

6.4.3.1 Comparison to health:

T-cell subsets of CCP+ individuals were categorised as normal if values fell within the 95% confidence limits of normal healthy controls.

Figure 29 represents the proportions of patients with abnormal T-cell subsets. If values in the CCP+ patients were similar to those of controls, then no more than 5% of patients would be expected to have values outside the reference limit. The dotted line denotes the expected proportion of abnormal values.

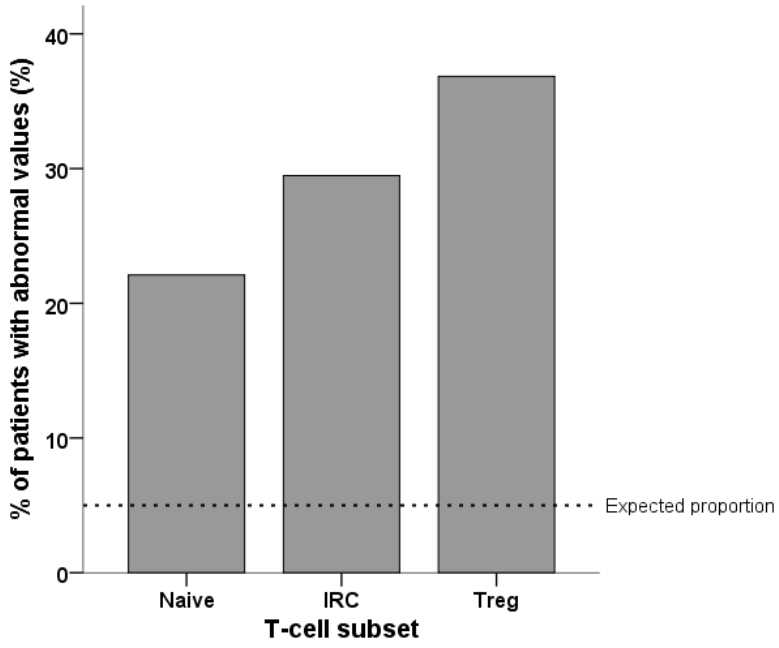


Figure 29 Proportions of patients with abnormal naive, IRC and Treg cell frequencies

Descriptively, CCP+ individuals demonstrated reduced naïve and Treg cells frequencies with elevated IRCs compared to health. In subjects with full T-cell data available (n=95/103), this variance from health was significant in each of the 3 T-cell subsets (Table 28, P<0.001 for all 3 subsets).

6.4.3.2 T-cell subsets in CCP+ individuals

Considering the frequency of abnormalities, over a third of CCP+ individuals had no T-cell abnormalities (37.9%), 40.0% had one, and the remaining subjects had two (18.9%) and three (3.2%) T-cell abnormalities (Table 28).

The figures below (Figure 30 & Figure 31) represents the proportion of patients progressing to IA according to (i) the T-cell subset status measured at baseline and (ii) the number of abnormal T-cell subsets. Data are presented as progression within 12 months and progression ever. The descriptive data suggest that a greater proportion of progressors have abnormal T-cell subsets compared to normal. Furthermore, those with the greatest dysregulation (2-3 abnormal T cell subsets) account for the largest proportion of progressors.

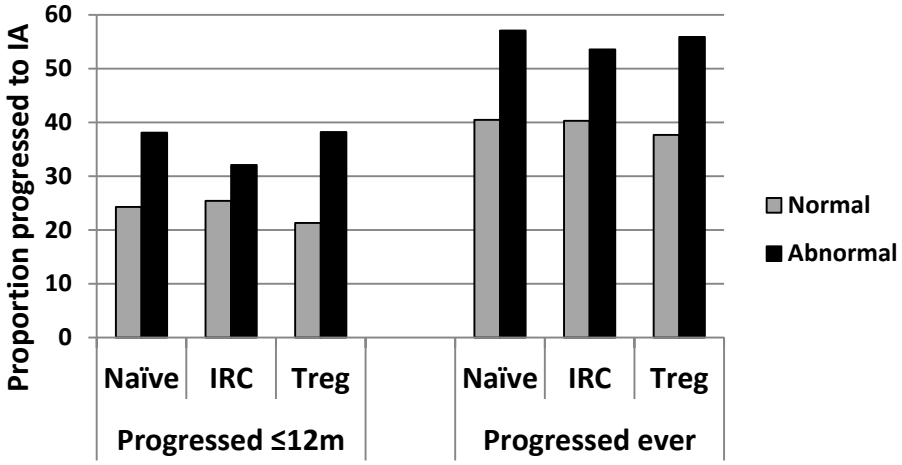


Figure 30 Proportions of patients progressing to IA within 12 months, or ever during follow-up, according their T-cell subset status at baseline

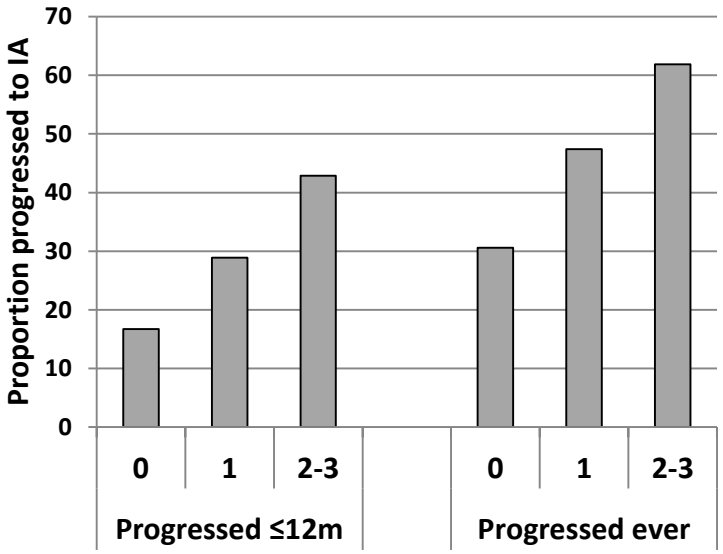


Figure 31 Proportions of patients progressing to IA within 12 months, or ever during follow-up, according the number of abnormal T-cell subsets at baseline

6.4.3.3 T-cell subset frequencies and progression to IA

In an unadjusted analysis, the association between T-cells and progression to IA was considered. The area under the receiver operator curve (AUC ROC) analysis demonstrated that T-cell subsets differed from 0.5 and were predictive individually, although weakly. Naïve AUC=0.63 (95% CI 0.52, 0.74), IRC AUC=0.63 (95% CI 0.52, 0.74) and Treg AUC=0.66 (95% CI 0.55, 0.77, all $p<0.03$). The best individual predictor was Treg, then naïve calls and lastly IRC (Table 28).

For each subset, the sensitivity and specificity for the eventual progression to inflammatory arthritis was calculated. Individually, naïve, IRC and Treg subsets demonstrated high specificities (Table 28, 71.7 – 83.0 %) for prediction of progression to IA, but had relatively low sensitivities (28.6 – 45.2%).

Consistent with the AUC ROC analysis, the unadjusted OR for all three subsets indicated that each subset was associated with progression to IA, where no adjustment was provided other than age (naïve and Treg) (Table 28). Higher naïve cell and Treg frequencies were protective (OR 0.94 95% CI 0.90, 0.98; OR 0.70 95% CI 0.56, 0.89 respectively), as would be expected. Higher IRC frequencies were associated with increased odds of progression (OR 1.15 95% CI1.00, 1.32), although the confidence interval includes one and therefore should be interpreted cautiously.

Table 28 Unadjusted T-cell analysis of progression to IA

	Reduced naïve cell frequency	**Elevated IRC frequency**	**Reduced Treg cell frequency**
Observed proportion of patients (observed/n)	22.5% 23/102	30.3% 30/99	35.4% 35/99
Calculated proportion†	22.1%	29.5%	35.8%
95% CI	14.2%, 31.8	20.6%, 39.7	26.2%, 46.3
standardised binomial test z	7.4	10.7	13.5
p	<0.001	<0.001	<0.001
AUC ROC †	0.63	0.63	0.66
95% CI	0.52,0.74	0.52,0.74	0.55,0.77
p	0.029	0.032	0.008
Sensitivity †	28.6 %	35.7 %	45.2 %
95% CI	17.2, 43.6	23.0, 50.8	31.2, 60.1
Specificity †	83 %	75.5 %	71.7 %

95% CI	70.8, 90.8	62.4, 85.1	58.4, 82.0
	Naïve (per %)*	**IRC (per %)**	**Treg (per %)***
Unadjusted OR †	0.94	1.15	0.70
95% CI	0.90, 0.98	1.00, 1.32	0.56, 0.89

*† in patients with data for all 3 T-cell subsets n=95, *Adjusted for age, AUC ROC=area under the ROC curve; OR=odds ratio*

6.4.3.4 Unadjusted analysis for time to IA

Time to development of IA from unadjusted data was achieved using Kaplan-Meier plots and log rank tests. These compared time to IA between those with and without abnormal T-cell subsets. Additionally, log rank tests were used to determine whether the total number of T-cell abnormalities (0, 1, 2-3) were related to time to progression to inflammatory arthritis (Table 29). Analysis was restricted to the 95 individuals with all 3 T-cell subsets.

Table 29 Median time to IA according to T-cell subset status (normal/abnormal) and number of abnormal T-cell subset frequencies present.

		Time to IA (months): median (95% CI)		
Subset	**Cut-off**	**Normal**	**Abnormal**	**Log-rank test**
Naïve	Less than LLN for age	50.1 (33.6, 66.7)	34.1 (0.0, 68.6)	X^2=2.2, p=0.142
IRC	>3.7	50.1 (40.2, 60.0)	22.1 (11.5, 32.6)	X^2=1.8, p=0.184
Treg	Less than LLN for age	46.5 (31.6, 61.3)	23.7 (1.1, 46.3)	X^2=2.3, p=0.126

	Time to IA (months): median (95% CI)			
Number of abnormal subsets	**0**	**1**	**2-3**	**Log-rank test**
	52.4 (36.7, 68.2)	44.4 (16.5, 72.2)	15.4 (0.0, 33.5)	X^2=5.54, p=0.062

Comparing each subset, individuals with abnormal subsets did correspond to a shorter time to inflammatory arthritis although not significantly. However, considering the number of abnormal subsets there was some indication that time to inflammatory arthritis did differ between patients with 0, 1 or ≥2 abnormal subsets (Chi-square=5.54, p=0.062).

Figure 32 A-C demonstrates the Kaplan-Meier plots associated to this analysis with time to IA dependent on each T-cell subset status. Figure 33 depicts the Kaplan-Meier plots when number of abnormal T-cell subsets are considered. In some of the plots the survival curves cross, suggesting violation of the proportional hazards assumptions. This was noted at the 48-month time point, when the number of patients still observed had reduced. This may affect the accuracy of the survival estimates beyond this point. It is reasonable to consider restricting analysis to progression to IA at 12 months given the imminent nature of the cohort.

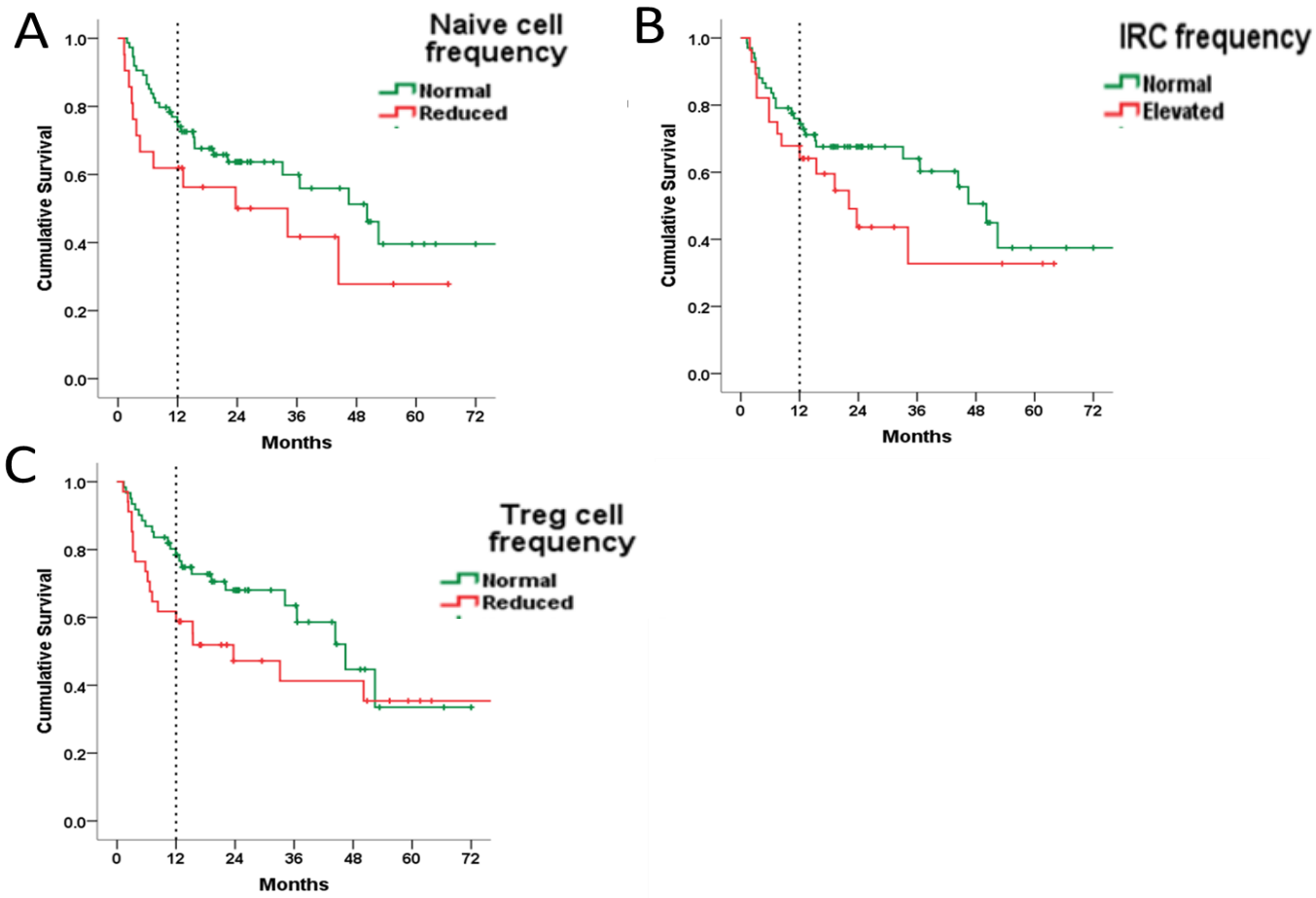


Figure 32 Kaplan-Meier plots of time to IA according to the status of each T-cell subset.

Figure 31A – Naïve cell, Figure 31B - IRC, Figure 31C - Treg. Green represents normal, represents abnormal, dotted line represent the 12 month follow up point

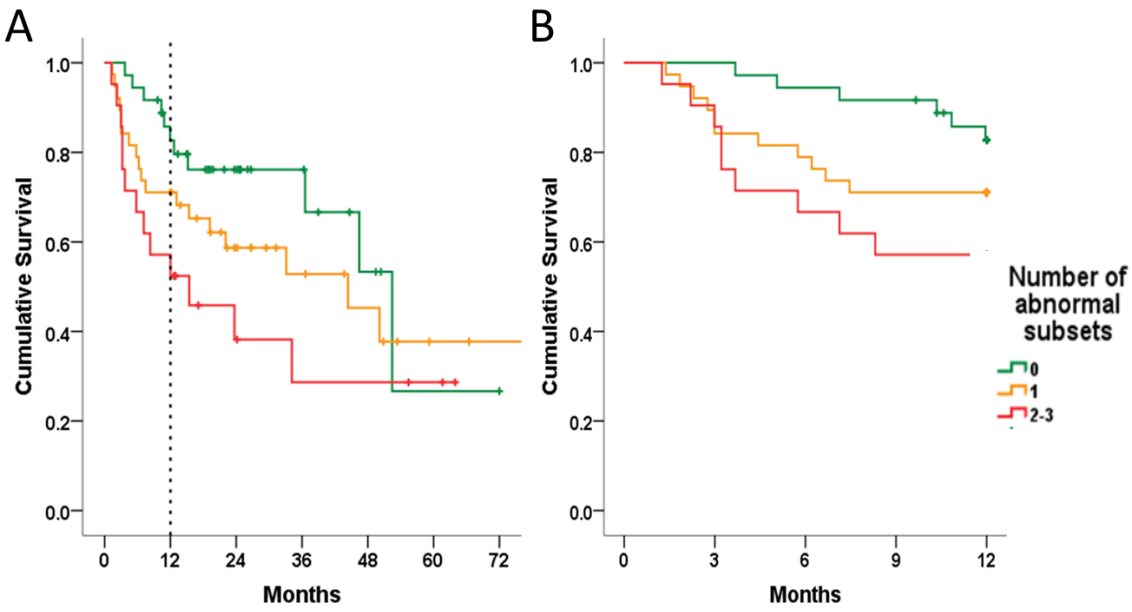


Figure 33 Kaplan-Meier plots of time to IA according to number of abnormal T- cell subsets.

Figure 32A - complete duration of follow up included. Figure 32B – restricted to only 12 months follow up. Green represents zero abnormal T-cell subsets, yellow represents one abnormal T-cell subset and red represents two to three abnormal t-cell subsets. Dotted line represents the 12 month follow up point.

6.4.4 Patterns of T-cell abnormality and associations with SE

The unadjusted analysis suggests an association between T-cells and progression. However, there are a number of confounding factors such as the presence of SE as well as environmental elements, in particular smoking [119, 408, 409]. A preliminary analysis of 90 individuals with both SE and T-cell data available was performed. Sixty-nine (72.2%) were found to be SE positive. There was no association at the 5% significance level, however descriptively a substantively greater proportion of SE positive patients had abnormally high IRC and low Treg frequencies with OR 2.51 (0.76, 8.23) and 2.25 (0.79, 6.38) respectively, (chi sq=2.39 p=0.122 for both). Furthermore, the frequency of abnormal T-cell subsets (out of 3) was substantially greater in SE positive individuals (Mann-Whitney U standardised test statistic z=1.47, p=0.143)

Table 30. The low number available for this analysis limits the significance of these findings. The descriptive data do suggest an association and certainly provide evidence that it should be considered in an adjusted analysis.

Table 30 Proportions of patients with 1, 2, or 3 abnormal subset values according to SE status.

	Number of T-cell subsets with abnormal values		
	1	2	3
Shared epitope			
Negative	48.0% (12/25)	8.0% (2/25)	- (0)
Positive	33.8% (22/65)	24.6% (16/65)	4.6% (3/65)

6.4.5 T-cell Model of Progression to IA

An adjusted multivariable analysis was required to ascertain whether there is value in measuring T-cell subsets over the clinical parameters routinely collected in clinic. These clinical parameters have previously been used to construct a prediction model ([217] and section 2.3).

In total, data from 103 patients were included in the multivariable models. Regression modelling was performed using T-cell subsets frequencies, controlling for age, SE, smoking status and clinical parameters, thus totalling nine predictor variables. Continuous T-cell data were used to retain as much information as possible.

Several intermediate models were developed to investigate the effect of confounding factors, such as genetic (SE) and environmental factors (smoking). Three primary models resulted for further consideration; a T-cell only model (model 1), a clinical only model (model 2) and a combined model (model 3), Table 31. In the combined model, some clinical parameters were removed since they were shown to be less predictive in intermediate models (see Model 3 for details). This allowed for a reduced number of variables given the limitation imposed by the relatively small samples size.

Model 1 (T-cell only)

When all three subsets were included in a model with age (Figure 34 and Table 31), naïve cell and Treg were independently associated with progression. The effect of IRC was less prominent. The AUC ROC for the predicted probability of progression from this model was 0.75 (95% CI 0.65-0.85), which represents an improvement over the prediction for all 3 individual subsets (Table 28).

Model 2 (clinical)

The clinical model consisted of antibody status (RF and/or CCP titre 3x the upper limit of normal), EMS >30 minutes and physician assessed small joint symptoms [217]. Within this particular group of CCP+ patients with arthralgia, EMS was not independently associated with the odds of progression to IA (Table 31, OR 1 95% CI 0.41, 2.42, p=0.997) discordant to the previously reported model. However, autoantibody status and the presence of small joint symptoms were still associated with progression (Table 31, OR 4.66 95% CI 1.21, 18.05, p=0.026 and OR 2.65 95% CI 1.14, 6.19, p=0.024 respectively). The AUC ROC for Model 2 was 0.62 (95% CI 0.54-0.76) Figure 34.

Model 3 (combined)

Adding the T-cell subsets to the clinical model was challenging due to the relatively small sample size for the number of variables assessed [407]. In such cases, it is recommended that the least significant of the variables in the full model are removed, provided this does not substantially affect the ORs for the remaining variables [410]. Considering the variables from model 1, model 2 and the confounders of SE and smoking; EMS (p=0.553) and smoking (p=0.627) were the least significant and were therefore removed. Removing smoking and EMS from the model did not substantially affect the ORs for the remaining variables. Age was retained (p=0.668) because its removal affected the ORs for naïve cell and Treg.

Following adjustment for age, SE, autoantibody status and joint counts, both naïve cell and Treg frequencies remained independently associated with the odds of progression (Table 31, OR 0.94 95% CI 0.89, 0.98, p=0.008 and OR 0.72 95% CI 0.55, 0.94, p=0.015 respectively), while the IRC association decreased (OR 1.05 95% CI 0.92, 1.20, p=0.441). The AUC ROC was 0.79 (95% CI 0.70-0.89) Figure 34. This is improved compared to models 1 and 2 showing the added value of combining both data sets.

Model Evaluation

In order to compare the performance of each model, AIC values were determined (lower values equate to better quality model). The combined model 3 represented an improvement over the clinical model 2 (Table 31, AIC 116.3 vs. 125.0). However, there was very little difference in the AIC between the combined model 3 and the T-cell alone model 1 (116.3 vs. 115.7). Despite this, the area under the ROC for the combined model 3 (0.79) was better than for model 1 (0.75). Importantly, both of these adjusted models improved on the values achieved by each subset individually (AUCs 0.63-0.66) and from the clinical only model 2 (AUC 0.62).

Table 31 Logistic regression models of progression to IA.

Logistic regression model				
		Model 1 (T-Cell)	**Model 2** (Clinical)	**Model 3** (Combined)
Naive (per %)*	OR	0.93		0.94
95% CI		0.89, 0.97		0.89, 0.98
p		0.002		0.008
IRC (per %)*	OR	1.07		1.05
95% CI		0.94, 1.23		0.92, 1.20
p		0.294		0.441
Treg (per %)*	OR	0.68		0.72
95% CI		0.53, 0.88		0.55, 0.94
p		0.003		0.015
Age (per year)	OR	1.01		1.01
95% CI		0.97, 1.05		0.97, 1.05

p		0.492		0.668
SE positive(%)	OR			2.36
95% CI				0.76, 7.36
p				0.138
Smoker				removed[†]
High positive RF /CCP[‡]	OR		4.66	2.79
95% CI			1.21, 18.05	0.58, 13.34
P			0.026	0.198
Small joint symptoms	OR		2.65	2.14
95% CI			1.14, 6.19	0.84, 5.46
p			0.024	0.110
EMS ≥30 mins	OR		1.00	
95% CI			0.41, 2.42	removed[†]
p			0.997	
AIC		115.7	125.0	116.3
AUC ROC		0.75	0.62	0.79
95 % CI		0.65, 0.86	0.54, 0.76	0.70, 0.89

** adjusted for age, AIC=Akaike information criterion; AUC ROC=area under the ROC curve,*

[‡]determined as >3xULN=upper limit of normal [†]removed from final model to reduce the number of covariates.

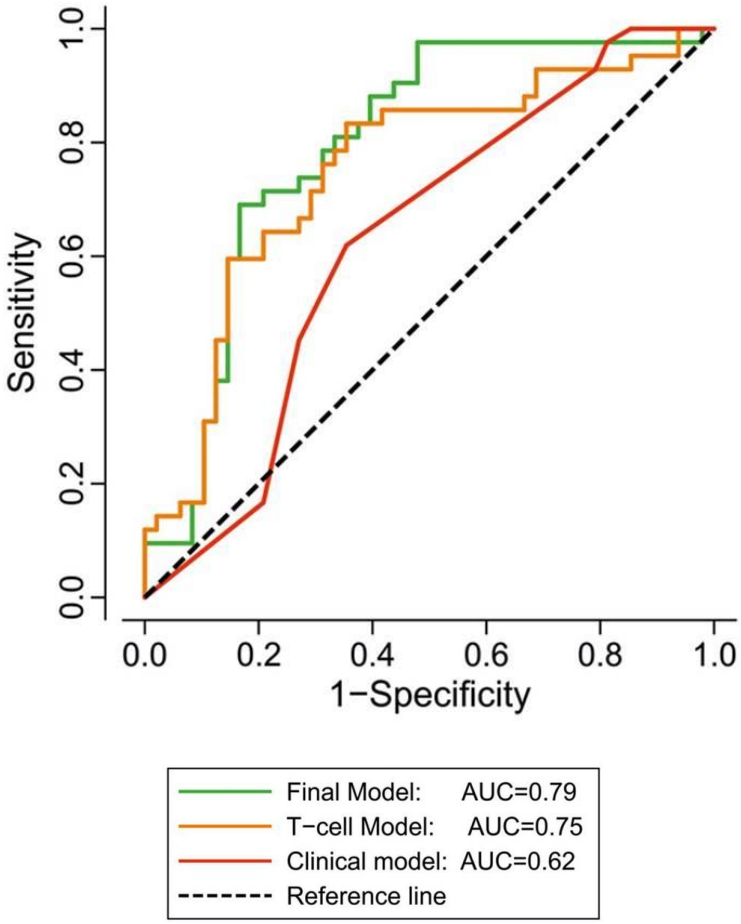


Figure 34 ROC graphical representation of the three logistic regression models.

Binary logistic regression models of the occurrence of progression to IA using model 1 T-cell subset only (orange line), model 2 clinical parameters (red line) and the combined model 3 (green line) were constructed. The area under the ROC for the predicted probability of progression from model 1 was 0.75 (95%CI=0.65-0.85), which represents an improvement over model 2 0.62 (95%CI=0.54-0.76). Model 3 showed the best results with area under the ROC at 0.79 (95%CI=0.79-0.89).

Model 3 was applied to the cohort enabling risk stratification into 3 categories according to their predicted risk of progression; low (0-19%, n=20), moderate (20-79%, n=56) or high (80-100%, n=14), Table 32. This followed the same method employed to assess the previously published clinical model [217]. A higher proportion of individuals progressed over time from the high-risk group (64%, 9/14) compared to those in other risk groups. The model also enables the low-risk group to be clearly identified with only 1/20 progressing ever.

Table 32 Proportions of patients progressing a) within 12 months or b) ever, according to their predicted probability of progression from model 3.

Predicted probability of progression	% progressed to IA within 12 months	% progressed to IA ever
Low (0-19%)	5% (1/20)	5% (1/20)
Moderate (20-79%)	38% (21/56)	57% (32/56)
High (80-100%)	29% (4/14)	64% (9/14)

6.4.6 Adjusted modelling for time to IA

Clinically, it is beneficial to provide patients with a time frame as to their likely progression. Therefore, similarly to the unadjusted analysis, a time to progression analysis was performed. Cox regression models were constructed using the three logistic regression models to investigate time to progression (Table 33). The trends identified were similar. All three T-cell subsets demonstrated association with the odds of progression in model 1 and 3 (Table 31). However, IRC were the most significant in this analysis. Using Harrell's C as an indication of performance of the Cox regression, the combined model 3 provided the best result. This allowed for 69% of randomly chosen pairs of progression times to be correctly ordered compared to 65% for the T-cell only model 1 and 60% in the clinical only model 2 (Table 33).

Table 33 Results of Cox regression models of time to progression to IA.

COX regression model		**Model 1 (T-cell)**	**Model 2 (Clinical)**	**Model 3 (Combined)**
Naive (per %)*	HR	0.97		0.97
95% CI		0.94, 0.99		0.95, 1.00
p		0.018		0.044
IRC (per %)	HR	1.08		1.08

95% CI		1.02, 1.15		1.01, 1.15
p		0.006		0.016
Treg (per %)*	HR	0.83		0.86
95% CI		0.70, 0.98		0.72, 1.02
p		0.027		0.091
Age (per year)	HR	1.01		1.00
95% CI		0.98, 1.03		0.98, 1.03
p		0.598		0.791
SE positive	HR			1.60
95% CI				0.66, 3.86
p				0.297
Smoker				not entered†
High positive RF / CCP‡	HR		2.44	1.45
95% CI			0.75, 7.92	0.41, 5.08
p			0.139	0.561
Small joint symptoms	HR		1.73	1.54
95% CI			0.96, 3.12	0.86, 2.77
p			0.071	0.149
EMS ≥30 mins	HR		1.21	
95% CI			0.66, 2.21	not entered†
p			0.536	
Harrell's C		0.65	0.60	0.69
AIC		329.6	335.7	332.4

** adjusted for age, AIC=Akaike information criterion; ‡ determined as >3xULN=upper limit of normal †removed from final model to reduce the number of covariates. HR=hazard ratio.*

As in the unadjusted analysis, CCP+ subjects were stratified into 3 groups according to their predicted risk of progression: low (0-19%) moderate (20-79%) high (80-100%) calculated using model 1, 2 & 3. Figure 35 presents Kaplan Meier plots for time to progression according to the predicted risk categories using each model. Time to progression differed significantly according to risk groups in both Model 1 (chi-sq=6.04, p=0.049) and model 3 (chi-sq=13.43, p=0.001), although there seemed to be little difference between the curves for patients at moderate or high-risk of

progression. In model 3 those within the high-risk group progressed to IA more rapidly (median 15.4 months, (95% CI=14.3-40.8)) compared to those in the moderate-risk (35.1 months (IQR 25.8-44.4)) and low-risk groups (63.4 months (IQR 57.9-69.3)). For model 2, no individual was categorised as high-risk. The median time to progression in the moderate risk group is 34.1 months with an overall significant difference between the two risk groups (chi-sq=4.60, p=0.032).

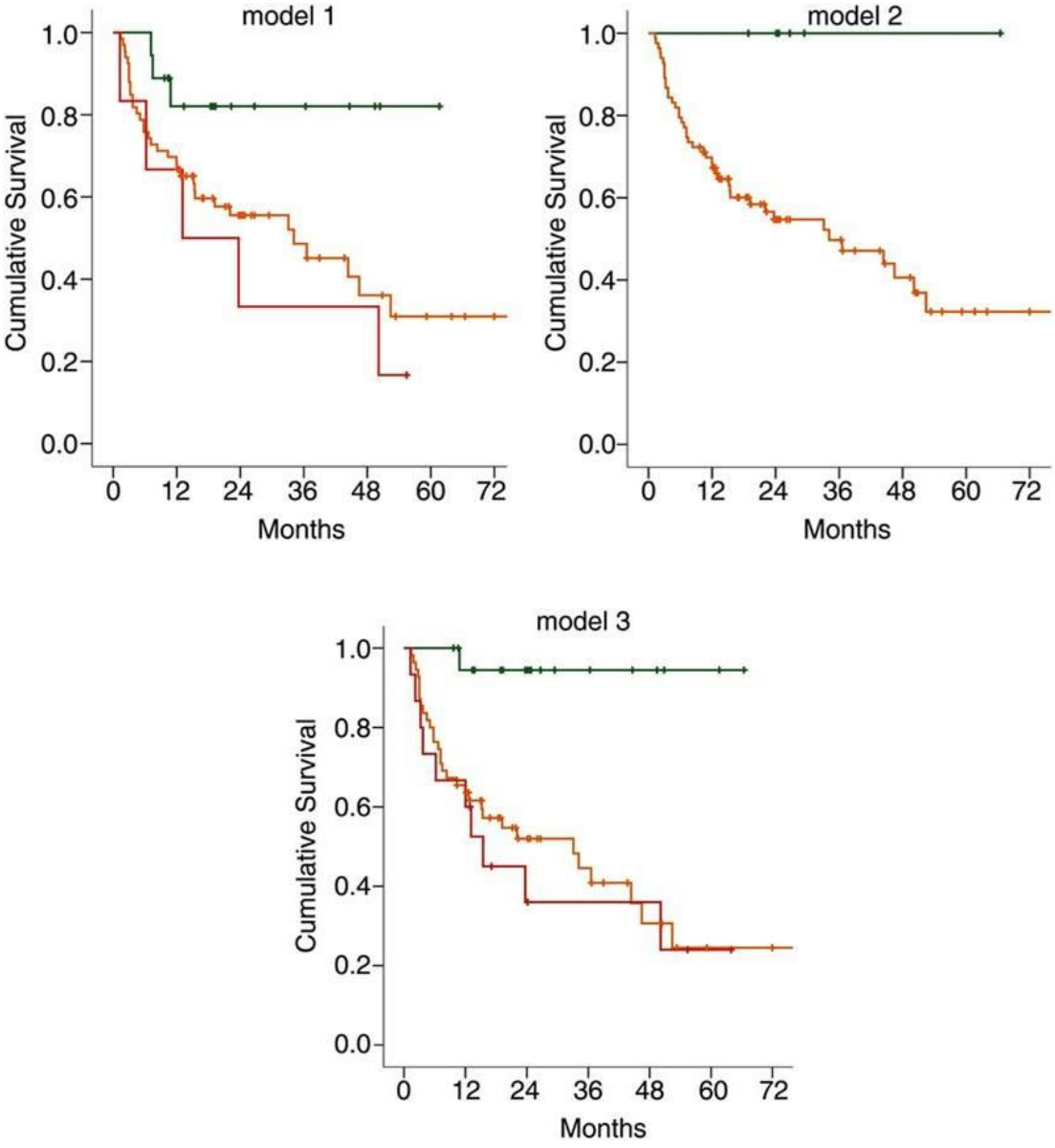


Figure 35 Kaplan-Meier graph of cumulative IA-free survival according to predicted probability of progression in the three models.

Kaplan-Meier plots for time to progression were constructed according to the predicted risk categories from logistic regression Model 1 2, & 3. Low risk (green line, 0-19%) moderate risk (orange line, 20-79%) high risk (red line, 80-100%).

6.5 Discussion

This is the first report exploring the clinical utility of T-cell subset quantification within a prospectively-followed 'at-risk' population. Given the paucity of data within this area it was first necessary to report on the pattern of T-cell subsets in CCP+ at-risk individuals compared to observations in health. As there was no existing reference range, a reference limit has been created. The development of reference limits permits the clinical utility of T-cell subsets as a biomarker. Using this reference limit, around two thirds of the CCP + at-risk cohort demonstrated some disturbance in frequency of T-cell subsets compared to health. Individuals at risk of IA had lower naïve and Treg cell frequency with elevated IRCs when compared to health. These differences were shown to be statistically significant and predictive of progression to IA individually, albeit weakly (Table 28). Although there is a lack of other studies which have considered T-cell subsets in at-risk individuals, the pattern of dysregulation reflects that seen in early RA with a reduction in naïve and Treg cells [247, 249, 250, 400].

These findings improve our understanding of the immuno-phenotype of these at-risk individuals. However, in order to ascertain the additional value of T-cell subsets in predicting progression, an analysis which accounted for the existing clinical model was required. The adjusted analysis reported an added value of combining 3 T-cell subsets and clinical data (Figure 34 and Table 31). The first step of this process demonstrated an added predictive value when combining the T-cell subsets into a model rather than using their individual value. This supports the notion that T-cells subsets could be considered together as a single entity. The second step which considered the previously established clinical variables (model 2) resulted in a lower area under the ROC compared to the T-cell model (model 1). Finally, in the combined model (model 3), T-cell subsets remained independently associated with the odds of progression even after incorporating the remaining clinical variables (joint count, CCP+ titre and adjusting for SE status). Area under the ROC was improved for this model when

compared to the T-cell and clinical models. However, when comparing models, additional measures such as the AIC is applied. The AIC is a measure that trades off the information a model provides about the outcome against its complexity. Analysing the AIC values, the final adjusted model (model 3) was similar to that of the T-cell model (model 1), although the AUC ROC was better for the combined model. It may be that validation of these models in a larger cohort is required in order to ascertain which model is optimal.

The survival analyses considering time to progression supported the hypothesis that those with the greatest T-cell subset dysregulation are at the highest risk of imminent progression (Figure 35, Table 33). This was suggested in the unadjusted analysis and further evidenced in the adjusted analysis. Here, all the models indicated that individuals who were categorised in the high risk group had a greater risk of imminent progression. The Kaplan-Meier curves for model-2 (T-cell only) consisted of only a low and moderate groups, as no individual was categorised in the high risk group. Since no individual in the low risk group progressed, a calculation for time to progression for that group was not possible. The combined model did allow categorisation into the high risk group, although there was some crossing of curves. Statistical significance however remained with the high risk group progressing earlier than the moderate risk group (15 months versus 35 months). This model therefore identifies those at risk of imminent progression. By establishing who is at risk of IA, clinicians can formulate better management plans and inform patients to be vigilant about change of symptoms.

There are limited studies to support or refute the findings reported here. No other research group has investigated T-cell subsets within a prospectively-followed at-risk population. A small cross-sectional study of 26 seropositive patients with arthralgia reported on peripheral naïve T-cells (CD3+CD4+CD45RO-CCR7+). However, no difference between health, RA and arthralgia was reported in this study [411]. It is worth noting that the antibodies used to measure naïve cells were not the same as has been reported in this work. Naïve CD4+ T-cells are conventionally identified using

markers such as CD45RA+, CD62L+ CD44+ but less commonly using an exclusion marker (CD45RO-) and an activation marker (CD197/CCR7) to more precisely define the subset. This makes a comparison of findings challenging. Furthermore, this group studied a mixed cohort of seropositive individuals with arthralgia (ACPA+ and/or IgM-RF+), compared to this study which recruited individuals based on CCP status alone. It is reasonable therefore to suggest there may be varying T-cell subset dysregulation between these cohorts.

Interestingly, when the descriptive data were reviewed (Figure 30) it is apparent that some individuals with normal T-cell subset frequencies still progressed to IA. This raises questions regarding the sequence of events leading to progression. T-cell disturbance and dysregulation are thought to be very early changes and it could be hypothesised that all individuals would demonstrate dysregulation prior to development of IA. One explanation for these findings may be related to the proximity of the sample acquisition and analysis from progression date. Dysregulation may occur subsequently, or just prior to progression and therefore a more recent analysis of T-cell subsets is necessary to exclude progression. Alternatively, it has been suggested that within the heterogeneous cohort of CCP+ at-risk individuals, not all will have (measurable) dysregulation in T-cell subsets and may exhibit a B-cell driven or alternative pathway of disease. This theory is corroborated by the finding that not all individuals with RA have abnormal T-cell subsets [399]. T-cell subsets shift and evolve with the influence of external factors and therefore patterns of dysregulation and normality change with disease progression and states of remission. Until the natural progression of T cell subsets in inflammatory disease is better understood these questions will be difficult to answer.

This work has focused on the biomarker potential of T-cell subsets and as yet there is not sufficient functional work or longitudinal data to comment further on the insights into disease onset and pathogenesis. Naïve and Tregs appear to be more important in the early stages of the IA continuum and in assessing progression risk. Whereas, IRCs are mostly driven by inflammation and therefore their value appears to be in

assessing time to imminent progression. This is particularly pertinent in at-risk individuals where routine inflammation markers are normal.

6.6 Limitations

As has been highlighted during the results and discussion section of this chapter, T-cell subsets are a dynamic biomarker which vary throughout an individual's progress along the inflammatory arthritis continuum. This analysis has used samples obtained from CCP+ (at-risk) individuals with arthralgia at their baseline appointment. Given the imminent nature of the changes seen in T-cell subset frequencies, it is likely that for individuals presenting early in their at-risk status, their immunophenotype has yet to become dysregulated. This resulted in an amendment to the protocol allowing for annual, sequential T-cell subset quantification. Data on T-cell subsets over time in an at-risk population will hopefully enable better understanding of the changes in each T-cell subset with progression or non-progression to IA. This data will be reviewed once sufficient sequential samples are available for meaningful analysis.

This analysis has not considered the ultrasound findings performed at baseline of the study. Baseline ultrasound scan reports were not available on a significant proportion of the individuals who had had T-cell subset quantification. This will be considered once the cohort has increased with sufficient follow-up and progression to IA observed. The lack of ultrasound data is acknowledged as a limitation in this current study.

Although this is one of the largest studies of CCP+ at-risk individuals and the first to consider T-cell subset dysregulation, the sample size has limited the robustness of statistical modelling. It is recommended that there should be at least 10 cases in the smallest outcome category ('events') per variable (EPV), although it has been shown that valid results can be obtained with EPVs between 5 and 9 provided the results are interpreted cautiously. In model 3 the EPV was 6.9, therefore these can be considered

as promising preliminary results, but this model must be considered exploratory until it is validated in a second cohort.

To date, flow cytometry in RA patients has been predominantly performed in a research setting, however during and following this project, development of a pathway within the local National Health Service (NHS) has been established. Several quality controls and audits have been performed to ensure replication of results between the clinical immunologists and the research scientists. A small proportion of the samples used in this analysis were performed in the clinical laboratories. Additional quality control checks were performed on these samples and review of the gating strategy applied by the NHS laboratories. However, it is acknowledged that during a time of transition there is the possibility that different technicians and laboratory equipment increase the opportunity for mis-reporting and errors. The additional checks put in place aim to reduce these occurrences. Despite the possible complications and difficulties during the transition period, this approach has demonstrated robust sample processing and importantly has significantly increased workload capacity compared to that within research laboratories. This increased capacity will facilitate the validation of these results.

6.7 Key points

- CCP+ individuals have a dysregulated T-cell phenotype with reduction in naïve and Treg proportions and elevation in IRCs when compared to healthy controls.
- T-cell subset quantification can be used for risk modelling to improve the prediction of progression to inflammatory arthritis.
- Individuals with the greatest immune dysregulation are at highest risk of imminent RA disease progression.

6.8 Conclusions

At-risk cohorts demonstrate great heterogeneity with regards to genetic predisposition, environmental factors, symptomatology, and ultimately progression towards arthritis. Although 53.4% (55/103) of subjects studied here have yet to develop IA, some progressed imminently - these individuals may benefit from early treatment. Identification of these individuals is therefore a clinical priority. Whilst the development of a clinical model has assisted this identification, the opportunity to improve prediction with the inclusion of immunological markers has been explored here.

This study has reported T-cell dysregulation within the CCP+ arthralgia cohort. The development of a reference limit enabled the dichotomisation of T-cells into normal and dysregulated. Clinical utility of T-cell markers has been explored through incorporation into regression modelling. Construction and comparison of models has demonstrated that the combined model (using both clinical and immunological variables) offered a predictive advantage over each variable when considered separately. Furthermore, the time to IA development was rapid in the high risk subgroup compared to the low risk subgroup. This was best appreciated in the models which included T-cell data.

To validate the findings of this study in a larger second cohort, further steps to ensure quality control measures for assessing T-cell subsets are required. Refinement of the reference range should be sought, particularly at the extremes of age. With ongoing recruitment to this prospective study the opportunity to include all biomarkers including ultrasound parameters is available. This may assist in establishing the value of immunological markers in predicting progression to IA. As indicated in the discussion, sequential samples will soon be available for these individuals and it would be beneficial to investigate the natural history of T-cell subsets following the development of IA. One would hypothesise that those individuals who do not progress to IA have relatively stable subsets. If so, this would provide an indication as

to whether progression is imminent and requires intervention or whether observation is appropriate. Adopting such a personalised medicine approach to services is an attractive option for both patients and physicians.

7. Concluding remarks and future directions

In this concluding chapter, a summation of results and implications of the research are presented. An update of the recent developments within the field is provided and future directions explored.

7.1 Summary and implications of findings

This thesis has focussed on an at-risk population defined by anti-CCP positivity and non-specific MSK symptoms recruited from the county of Yorkshire, U.K. The previously presented chapters report on the findings of imaging, molecular and cellular markers which contribute to our understanding of disease progression and prediction of IA development.

Magnetic Resonance Imaging in individuals with systemic autoimmunity and arthralgia: MRI as an imaging biomarker

For the first time, the MRI characteristics of a large cohort of CCP+ at-risk individuals have been reported. Within this populations, tenosynovitis was found to be the most prevalent finding with comparatively low levels of BME and erosions. Previous work concerning the clinical and ultrasound variables facilitated multivariable modelling to evaluate the ability of MRI features, including tenosynovitis, to predict IA. MRI tenosynovitis was predictive of future IA development in at-risk individuals and predicted clinical synovitis in an individual joint. In addition, BME was significantly associated and predictive of clinical synovitis development in a joint although did not predict progression at a patient level. The study has confirmed that MRI variables, in particular MRI tenosynovitis, provide additional predictive ability over and above the variables included in the clinical model and ultrasound GS & PD. Whether tenosynovitis and BME are the primary initiating lesions prior to IA development has not been confirmed by this analysis, but can be the focus of future investigation.

The results from this study may have significant implications for how at-risk individuals are assessed. Although a preliminary report, the high HRs for MRI tenosynovitis suggest superiority with improved prediction compared to ultrasound in this at-risk cohort. A comprehensive analysis comparing the detection of tenosynovitis by ultrasound and MRI is required to ascertain whether the modalities are comparable. Currently, individuals recruited to the study have both ultrasound and MRI. However, outside of the research setting, it may be conceivable that one imaging examination is sufficient. MRI offers an alternative and comparable, arguably superior option that may be more feasible, where ultrasound expertise is unavailable.

MicroRNA profiling of matched samples in individuals with systemic autoimmunity and arthralgia who progress to RA

In this exploratory study, miRNA global expression profiling identified several miRNAs of interest that were dysregulated with progression. In contrast, to previous research, this project defined the study population further, by focussing on those individuals with no evidence of inflammation as determined by PD on ultrasound. To date, there has been no other study which has reported on predictors of disease in such an early phase of autoantibody positivity. At this time, an exact miRNA signature cannot be validated. Instead there are several miRNAs identified which are expressed during the CCP+ state and on through to progression. A multivariable analysis to evaluate prediction to progression was restricted given the sample size. However, the potential utility of miR-22 as a biomarker to differentiate at-risk individuals was highlighted and through further validation may offer clinical utility. The scientific value of the study design using sequential samples prior to and at point of progression should not be dismissed. Such a design offers the opportunity for work evaluating the mechanistic functions of specific miRNAs.

From a scientific standpoint, understanding the dysregulation in molecular markers provides insights into disease, although does not necessarily translate into clinical utility. The identification of molecular biomarkers measurable from serum samples in

the validation phase, has significant potential implications. Through further analyses and validation, it may be possible to improve prediction models or develop a pathway in which a blood test is sufficient to establish risk. Furthermore, this is the first study to report on an at-risk population in which subclinical inflammation was excluded. The fact that individuals still progress, exemplifies the need to identify those at greatest risk. Whilst conducting this research, inconsistences with published methodology and analytical approaches in miRNA studies became evident. This impedes the ability to compare findings and highlights the importance of validation in different populations to establish the full extent of associations. Future work may look to improve upon methodology used and move towards collaborative projects to reduce inconsistencies.

T-cell subsets in individuals with systemic autoimmunity and arthralgia: an immunological biomarker

The development of a healthy control reference range permitted the evaluation of T-cell subsets as a novel biomarker in the prediction of IA progression within an at-risk cohort. T-cell dysregulation was evident with the extent or depth of dysregulation appearing to be greatest prior to IA development. Through quantification of T-cell subsets, prediction of IA was possible and superior to using the clinical model alone. A combined prediction model consisting of clinical and cellular markers demonstrated the greatest clinical utility. The effect of ultrasound and other imaging markers has not been evaluated in this study and should be a focus for future work (section 7.3).

Incorporation of cellular immune markers to the clinical model has demonstrated improved prediction and risk stratification. This has implications for the assessment of individuals and also generates questions concerning the sequence of events leading to IA development. Determining the earliest detectable change or patterns indicating stabilisation may assist identification of an optimal time point for intervention. This study has also demonstrated how biomarkers identified through research can be

successfully transferred to a local NHS setting, improving the access of T-cell subsets quantification.

7.2 Developments within the area of interest

Developments in defining at-risk:

Recruitment and investigation of several at-risk cohorts continues to provide further insights on the associations of disease progression. The clinical symptoms and initial presentation of individuals at risk of RA development has been the focus of extensive qualitative research [412, 413]. Individuals with seropositive arthralgia have reported experiencing a multitude of symptoms akin to those with newly diagnosed RA. A broad range of symptoms were reported including pain, stiffness, fatigue, burning sensations, loss of motor control, weakness, muscle cramps, sleep disturbance and abnormal skin sensations. Such information can help in understanding health seeking behaviours in such individuals, however it is limited to the defined ACPA positive population. Interest in identifying individuals across all phases of risk (without limiting to those with positive serology), has resulted in a definition of the prodromal symptom phase prior to RA development. Through assimilation of the available data, consensus of opinions and a validation exercise, the EULAR taskforce proposed 7 characteristics, listed in Box 6 [414]. This was designed to assist physicians in identifying individuals with clinical CSA prior to any laboratory or imaging investigations. Its role in primary care setting has not been suggested or evaluated. A high sensitivity (>90%) is obtained if ≥3 parameters are present, and a specificity (>90%) requires ≥4. It has been validated in other large cohorts and will hopefully enable future work to develop a criteria for imminent RA [415].

Box 6 EULAR defined characteristics describing arthralgia at risk for RA

History taking:
Joint symptoms of recent onset (duration <1 year)
Symptoms located in MCP joints

Duration of morning stiffness ≥60 min
Most severe symptoms present in the early morning
Presence of a first-degree relative with RA
Physical examination:
Difficulty with making a fist
Positive squeeze test of MCP joints

Adapted from van Steenbergen et al [414]

Note: These parameters are to be used in patients with arthralgia without clinical arthritis and without other diagnosis or other explanation for the arthralgia

As the field moves towards prevention and interventional studies, investigators have conducted qualitative research to ascertain perceptions around risk. Interviews and focus group discussions in both autoantibody positive individuals and FDRs have demonstrated key themes concerning the psychological implications of an at-risk status diagnosis [412, 413, 416, 417]. Within autoantibody positive individuals, the implications of symptoms on an individual's well-being have been reported [412]. Themes emerged regarding fear of future RA diagnosis, uncertainty, frustration and despair [413]. Acknowledging that an at risk of RA diagnosis may induce some psychological distress suggests care pathways need to be refined to assist in the management or indeed prevention of such conditions. By appropriately risk stratifying and identifying those at low risk, management can be tailored to develop reassurance and educational resources to help understand their symptoms. Similar studies in FDRs have also indicated anxiety around future diagnosis of RA [416, 417]. Concerns were raised as to the negative implications of predictive testing, although many did recognise how research in the area assists development and implementation of preventive strategies [417]. These studies concluded that communication of risk needs to be appropriate and address the psychological burden associated with predictive testing and involvement in studies.

Currently, limited data are available with regards to perceptions concerning interventional therapy in these cohorts. Interviews with FDRs have reported a willingness to receive 'prophylactic' treatment, which the majority of participants considered would be medication [418]. There was also a willingness to participate in RCTs to prevent or delay disease. A larger study analysing survey responses from 32 FDRs was able to demonstrate an increased odds ratio for likelihood of taking treatment if a risk reduction >20% of RA development was demonstrable, with a low probability of adverse events of therapy (<10%) [419]. Interventional studies should consider the acceptability of preventative therapies as these can vary depending on the patient population, communication of risk and duration, mode and safety of therapy offered.

Developments concerning interventional studies:

Preliminary results from the PRAIRI study have been presented [420]. This multicentre, randomised, double-blind, placebo-controlled clinical trial successfully recruited 82 autoantibody positive subjects. The objective was to assess whether a single infusion of rituximab could delay or indeed prevent the onset of clinical arthritis. The inclusion criteria stipulated positivity for both ACPA and RF as well as CRP levels ≥3 mg/l and/or subclinical synovitis on ultrasound or MRI of the hands. Each group received premedication 100mg methylprednisolone. However, given the findings from the intramuscular steroid study this would not be expected to affect onset of arthritis (section 2.4.1 and [293]). Eighty-one subjects were randomised and were observed for a median of 27 months (IQR 25 month). Thirty subjects developed arthritis during the observation period: 16/40 (40%) in the placebo group and 14/41 (34%) in the rituximab group, after a median period of 11.5 (IQR 12.5) months in the placebo group versus 16.5 (IQR 19.0) months in the rituximab group. In the preliminary analysis time to arthritis was significantly delayed in the rituximab compared to the placebo treated group (p<0.0001). The authors acknowledge that further follow-up will be significant in clarifying impact of a single infusion of rituximab.

This study is the first to consider biological DMARDs in an at-risk population and has generated extensive discussion. There are several areas for consideration when evaluating the significance of this study. Firstly, only one dose of therapy was administered in this trial. It is plausible to hypothesise that a longer duration of therapy may have led to further delay and potentially prevention. Furthermore, the inclusion criteria required individuals to have either elevated inflammatory markers or imaging findings of subclinical inflammation. From the findings of the Leeds cohort, it is recognised that individuals with subclinical inflammation (as determined by ultrasound and MRI) are a greater risk of IA development and as the goal posts shift may even be considered as early RA. It is therefore, not surprising that early intervention delayed but did not prevent clinical disease presentation. As a first step this study has been beneficial in demonstrating delay in clinically evident disease progression, however it could be viewed as successful therapy with relapse rather than prevention. Understanding the benefit of therapy in earlier phases of at risk is also required. Whether individuals at lower risk as defined by the absence of clinical imaging synovitis would benefit from interventional therapy is not known. There is a possibility for potential side effect of treatment outweighing any risk reduction benefit. Certainly, evaluating a moderate or lower risk group would require longer observational periods as the time to onset of disease may be longer.

A UK based study (Arthritis Prevention in the Pre-Clinical Phase of RA with Abatacept-APIPPRA) is hoping to add further evidence as to the implications of targeted intervention in at-risk individuals with the use of abatacept (2.4.3 and [421]). The protocol in this study is for 52 weeks of abatacept therapy. Additionally, the exclusion criteria include subclinical inflammation as determined by ultrasound findings. This longer duration of therapy and in a different population of at-risk individuals, potentially at lower risk, offers the opportunity to target disease at the earliest phase. Recruitment continues and preliminary results are eagerly awaited.

The inclusion and exclusion criteria of these studies have enabled a defined phenotype of at-risk individuals to be evaluated. However, as yet, an interventional study that

risk stratifies eligible individuals for the purpose of therapeutic management has not been described. Such study protocols of stratified biomarker RCTs have been pioneered in oncology and enable several biomarkers and therapies to be evaluated in a biomarker enriched population [422, 423]. In the context of individuals at risk of RA, use of biomarkers followed by risk stratification would allow for targeted therapy to those at greatest risk. Additionally, those at low risk may be offered health promotion and symptom management education. Given the dynamic nature of these cohorts, it is important that individuals are re-evaluated throughout the observation period since some individuals will change risk status. Adopting such a personalised or tailored approach to patient care should provide the optimum outcomes without exposing individuals at low risk to therapy side effects.

Developments in primary prevention:

In the U.K, a national registry of FDRs has been established [424]. Following work conducted in similar cohorts in the U.S, this registry aims to consider the interactions between the environment, genes and development of immune and inflammatory responses. A sub-study of the registry will also consider cardiovascular risk factors of disease. As yet, no data have been published and ongoing recruitment is required for sufficient numbers to enable a meaningful analysis. It is hoped that research findings will inform researchers and clinicians as to which FDRs are at greatest risk of RA and possible identification for intervention strategies.

Lifestyle modification may provide the opportunity to achieve true primary prevention in individuals at-risk. In support of this approach, qualitative research in FDRs has suggested that preferred risk reduction strategies would consist of lifestyle modification measures as opposed to medication [416]. In the U.S. a RCT entitled 'Personalized Risk Estimator for Rheumatoid Arthritis (PRE-RA) Family Study', is testing the effects of personalised risk education including modifiable factors such as oral health, diet, weight and smoking status in FDRs [425]. Results have recently been published and demonstrate that individuals in the intervention arm were statistically

more likely to demonstrate changing behaviours than those receiving standard RA education [426]. This study has provided evidence that educational tools personalised to an individual can increase motivation to alter behaviour. It is hoped that such intervention will prevent RA progression, as well as the other significant health benefits.

7.3 Future directions

Through the summation of the results presented in this thesis, it has been possible to refine existing models of prediction and provide insights into potential new markers. There are however, several areas to direct future work.

The MRI study has highlighted the improved prediction of RA development offered by MRI tenosynovitis in seropositive (CCP+) individuals. Preliminary results comparing both MRI and ultrasound tenosynovitis reporting, suggested that MRI was more sensitive. Limitations, as previously stated, inhibited a comprehensive comparison. Through validation in a larger cohort the associations with ultrasound tenosynovitis should be clarified. This would enable researchers to provide recommendations as to which imaging modality is superior in evaluating individuals at risk and negate the need for unnecessary investigations.

An area of interest that has arisen from this work relates to analysis of the MRI dynamic sequences. OMERACT RAMRIS scoring was used in this analysis based on its reproducibility and transferability for multicentre use [427]. Within this study, MRI synovitis was not associated with development of future clinical synovitis in a joint. Synovitis, as determined by RAMRIS scoring, is an ordinal measure that does not give any indication of severity or intensity of synovitis. Alternative assessments which provide continuous measures are available and are suggested to be more sensitive [428]. Such quantitative scores include synovial volume and dynamic contrast enhancement (DCE) MRI which have been successfully used [429].

Synovial volume is a good marker of disease activity however, the process of manually outlining synovitis is laborious. There are now automated and semi-automated systems available which provide a feasible option for larger scale studies [430, 431]. DCE-MRI offers assessment of the uptake and washout of gadolinium based contrasts and provides data on the synovial perfusion and capillary permeability [429, 432, 433]. Difficulties ensuring the consistency in characterisation of the enhancement curves and defining the regions of interest have limited its widespread use. However, it has been demonstrated that DCE-MRI can be used in multicentre trials and with good repeatability [434]. In RA cohorts, DCE-MRI has been shown to correlate with clinical features including joint swelling, pain and disease activity scores [435-437] and differentiate between active versus inactive synovial change [435]. DCE-MRI has also been shown to correlate with US PD [438, 439]. This may be particularly pertinent to at-risk cohorts since active synovitis may theoretically be associated with a progressive outcome. Furthermore, the value of DCE-MRI may be particularly useful for evaluating change in synovitis across longitudinal scans.

The miRNA and T-cell subset studies have illustrated how novel biomarkers can be identified and applied to a clinical cohort. Whilst, the miRNA work is preliminary, the possibility of a signature of miRNAs to assist in evaluating individuals at the very earliest phase of at-risk (without ultrasound powerDoppler) has been explored. It would be important to report on whether this miRNA dysregulation is also seen in those individuals with inflammation at baseline imaging. A correlation between miRNA dysregulation and ultrasound or MRI determined tenosynovitis has not been performed but should form the basis of future work. Similarly, association with T-cell subset and ultrasound findings has not been reported and is an area that requires consideration.

Other T-cell subtypes have been proposed in RA pathogenesis, including Th17 cells. Since the discovery of the CD4 effector T helper cell which produces interleukin 17 (IL-

17), the focus of many studies has evolved to consider the relationship between Th17 and disease [440-443]. The pro-inflammatory properties of IL-17 as well as its stimulation of osteoclastogenesis and degradation of cartilage has implicated Th17 as the driving force for autoimmunity [444-446]. However, there have been discrepancies in the surrogate markers used for Th17 identification, which may have contributed to the inconsistencies in results presented. Th17 is detectable in early disease, although levels are thought to reduce with therapy and disease chronicity [443]. Prior to the development of arthritis, IL-17 is markedly elevated, with a reduction at disease onset [196]. Within a seropositive arthralgia cohort (n=26), Th17 lineage cells (CD4+CD161+) were reported to be elevated [411]. The quantification of Th17 alongside naïve, Treg and IRC would provide a comprehensive immunological picture of at-risk individuals. Future work aims to address this. Furthermore, through the development of a reference range for Th17, the clinical utility can be assessed in prediction models.

Ultimately, there remains no single biomarker or 'magic bullet' that can predict progression in these individuals to date. Instead, risk stratification is possible through the integration of multiple cellular, molecular, imaging and clinical markers. As a consequence of relatively modest numbers of individuals in at-risk cohorts, the evaluation of all markers in a hazard regression analysis has currently not been performed. Future work requires such an analysis in order to determine whether progression can be predicted and which markers provide greatest information. With ongoing recruitment, the opportunity to include all biomarkers in an analysis to refine and adapt the models should be achievable.

Thus far, studies in this cohort, have reported on the predictors of progression to enable early identification of those at greatest risk [217, 237, 382]. However, the prospective design of this study enables data collection at multiple time points. There is the opportunity therefore, to assess disease evolution and natural history, from at risk with autoimmunity to RA development. Evaluation of imaging, cellular, molecular and clinical variables could provide evidence as to the sequence of events leading to

persistent joint inflammation and disease. It is apparent from the risk stratification exercises conducted to date, that predicting development of RA from baseline characteristics is not yet fully achievable. These individuals, by virtue of being part of a dynamic continuum, will inherently evolve and change phenotype. The miRNA project illustrates the concept of following an individual from presentation to progression and evaluating change in miRNAs expression. Similarly, several individuals within the cohort had sequential T-cell subsets quantified during follow-up. Preliminary data has been reviewed descriptively to determine stability of T-cell subset and whether there is change in dichotomisation groups with time to RA development. Analysis has been limited to a subset of patients and typically focussed on one characteristic or variable. An overview of changes in all variables is therefore warranted.

Obtaining data on changing phenotypes with RA development will assist in developing effective care pathways. It is hypothesised that immunological markers would be one of the first indicators of change followed by imaging and finally, the reported clinical symptoms. This analysis will require a near complete dataset with imputation to account for missing variables. Initially, it is proposed that latent growth curve models are used to evaluate change in characteristics of the progressors. This can then be compared to analysis within the non-progressors. By establishing the natural history of change, individuals can be re-evaluated and stratified throughout their care pathway. A personalised approach is proposed in which an individual can be identified as their markers evolve to represent a higher risk phenotype, providing the opportunity for communication of risk and possible management options.

As illustrated by the health promotion study in the U.S, aiming for primary prevention is a potentially achievable goal in at-risk cohorts [426]. Establishing the contributors and triggers to disease has resulted in studies considering the first sites of possible antigen presentation such as the mucosa (Section 2.2.3.5). Within our department, studies are underway to determine the associations between periodontal disease and RA progression through comprehensive periodontal examinations and mucosal

sampling in the CCP positive, at-risk cohort [283]. Similar work is planned for the assessment of autoimmunity at the site of the lungs through induced sputum analyses. However, it is recognised that in studying individuals with systemic autoimmunity the prospect of primary prevention is limited. Future plans therefore are to offer similar assessments to FDRs who are known to have genetic susceptibility, but as yet not developed systemic autoimmunity. By determining the significance of mucosal autoimmunity prior to systemic, there is the opportunity to intervene. These are currently pilot designed studies however, depending on their findings, could prove an important focus for future work.

It is apparent from the summation of findings and work proposed that there remain several areas which require consideration in the evaluation of risk in seropositive individuals. A larger validation process is required to establish the appropriate biomarkers to be used for successful risk stratification of at-risk populations. Secondly, little is known as to the evolving nature of RA development in these individuals and therefore a programme of work considering change in biomarkers and characteristics has been described. Since preliminary work has indicated a progression rate of around 50% in these individuals, identifying those at imminent risk remains a priority. The heterogeneity seen within a CCP positive at-risk status illustrates the need to address those at moderate to low risk whom may change phenotype. Similarly, there remain those considered at very low risk of progression whom should be appropriately reassured to minimise on psychological implications. To address these needs, a risk stratified interventional study is proposed. Drawing upon what has been established in at-risk cohorts, a study design which successfully risk stratifies individuals with appropriate communication of risk to facilitate recruitment is suggested. Whilst individuals with systemic autoimmunity remain a target population, FDRs and seronegative individuals with CSA features should be considered in studies moving forward. The proposed work should provide evidence as to when the window of opportunity truly commences and if indeed intervention in these early phases can interrupt or prevent disease onset.

List of Abbreviations

ACPA	Anti citrullinated protein antibody
ACR	American college of rheumatology
AIC	Akaike information criteria
Anti-CCP	Anti cyclic citrullinated peptide antibody
Anti–Car P	Anti carbamylated protein antibody
APC	Antigen presenting cell
Apo A1/B	Apolipoprotein A1/B
AUC ROC	Area under the receiver operator curve
BMC	Base of metacarpal
BMI	Body mass index
BME	Bone marrow oedema
CI	Confidence interval
CRF	Case report form
CS	Clinical synovitis
CVD	Cardiovascular disease
DCE	Dynamic contrast enhance-d/-ment
DMARD	Disease modifying anti-rheumatic medication
DRadius	Distal radius
DRUJ	Distal radioulnar joint
DTT	Dithiothreitol
DUlna	Distal ulna
EMS	Early morning stiffness
EQ-5d	Euroquol- 5d questionnaire
ESR	Erythrocyte sedimentation rate

EULAR	European league against rheumatism
EPV	Events per variable
FCR	Flexor carpi radialis
FCU	Flexor carpi ulnaris
FDR	First degree relative
FDSFDP	Flexor digitorium superficialis and flexor digitorium profundus
FPL	Flexor pollicis longus
FLS	Fibroblast-like synoviocytes
GWAS	Genome wide association studies
HAQ	Health assessment questionnaire
HC	Healthy controls
HDLc	High-density lipoprotein cholesterol
HLA	Human leucocyte antigen
HR	Hazard ratio
HsCRP	High-sensitivity CRP
IA	Inflammatory arthritis
ICJ	Intercarpal joint
INF-	Interferon-
IQ	Interquartile range
IRC	Inflammation related cells
IL -	Interleukin -
ILC	Innate lymphoid cell
IRF	Interferon regulatory factor 5 gene
LDLc	Low-density lipoprotein cholesterol
MCP	Metacarpophalangeal

MiRNA	Micro ribonucleic acid
mRNA	Messenger ribonucleic acid
MR	Magnetic resonance
MRI	Magnetic resonance imaging
MSK	Musculoskeletal
MTP	Metatarsal phalangeal
NHS	National health service
NPV	Negative predictive value
OMERACT	Outcome measures in rheumatology
OR	Odds ratio
PAD	Peptidyl arginine deiminase
PB	Peripheral blood
PBMCs	Peripheral blood mononuclear cells
PCT	Procalcitonin
PD	power Doppler
PET	Positron emission tomography
PIP	Proximal interphalangeal
PPV	Positive predictive value
PTPN-22	Protein tyrosine phosphate non-receptor-22
PSS	Primary sjogrens syndrome
qPCR	Quantitative polymerase chain reaction
RA	Rheumatoid arthritis
RAMRIS	Rheumatoid arthritis MRI Score
RASF	Rheumatoid arthritis synovial fibroblasts
RCJ	Radiocarpal joint

RF	Rheumatoid factor
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
ROC	Receiver operator curve
SD	Standard deviation
SE	Shared epitope
SF	Synovial fluid
SNP	Single nucleotide polymorphism
SPLA2	Secretory phospholipase A2
ST	synovial tissue
TC	Total cholesterol
TG	Triglycerides
TLDA	TaqMan low density arrays
TNF-alpha	Tumour necrosis factor-alpha
Treg	T-regulatory cells
UA	Undifferentiated arthritis
USS	Ultrasound scan
US	Ultrasound
VAS	Visual analogue score
VERA	Very early RA

Appendix A

ULTRASOUND JOINTS

R/L HANDED:

VISIT: Week 0 / 26 / 52 / annual (after week 52) / unscheduled visit / withdrawal

DATE:

RIGHT						LEFT				
EC*	GE (0-3)	#E (0-3)	PD (0-3)	GS (0-3)		GS (0-3)	PD (0-3)	E (0-3)	GE (0-3)	EC*
					ICJ					
					UCJ					
					RCJ					
					GLOBAL WRIST					
					MCP1					
					MCP2					
					MCP3					
					MCP4					
					MCP5					
					PIP1					
					PIP2					
					PIP3					
					PIP4					
					PIP5					
					SHOULDER					
					ELBOW					
					KNEE					
					ANKLE					

					MIDFOOT					
					1 MTP					
					2 MTP					
					3 MTP					
					4 MTP					
					5 MTP					
					OTHER					
					OTHER					
					OTHER					

ULTRASOUND SCAN: TENDONS

RIGHT			LEFT	
PD (0-3)	GS (0-3)		GS (0-3)	PD (0-3)
		ECU		
		MCP2_FT		
		MCP3_FT		
		MCP4_FT		
		MCP5_FT		
		PIP2_FT		
		PIP3_FT		
		PIP4_FT		
		PIP5_FT		
		OTHER		
		OTHER		
		OTHER		

*EC = Extra-capsular abnormalities. If present enter the appropriate number.

Appendix B

Score sheet for the OMERACT RAMRIS

using the EULAR-OMERACT RA MRI reference image atlas

WRIST JOINTS

MRI ID: _____ **Scorer's name:** _____

Centre where MRI was performed: _____

Image set (e.g. baseline or follow-up): _____

Sequences scored: _____

Scoring of synovitis

	Distal radio-ulnar joint	Radio-carpal joint	Intercarpal-CMCJ
Synovitis (0-3)			

Scoring of bone erosion and bone oedema

Bone erosion is scored 0-10, according to the proportion (in increments of 10%) of bone involved: 0: 0%, 1: 1-10%, 2: 11-20 %, ……., 10: 91-100%

Bone oedema is scored 0-3, according to the proportion (in increments of 33%) of bone involved: 0: 0%, 1: 1-33%, 2: 34-66 %, 3: 67-100%

For carpal bones, score the whole bone. For long bones, score from the articular surface (or its best estimated position if absent) to a depth of 1 cm.

	Base of metacarpal				
	1	2	3	4	5
Bone erosion (0-10)					
Bone oedema (0-3)					

	Trapezium	Trapezoid	Capitate	Hamate
Bone erosion (0-10)				
Bone oedema (0-3)				

	Scaphoid	Lunate	Triquetrum	Pisiform
Bone erosion (0-10)				
Bone oedema (0-3)				

	Distal radius	Distal ulna
Bone erosion (0-10)		
Bone oedema (0-3)		

Score sheet for the OMERACT RAMRIS

using the EULAR-OMERACT RA MRI reference image atlas

MCP JOINTS

MRI ID: _____Scorer's name: _____

Centre where MRI was performed: _____

Image set (e.g. baseline or follow-up): _____

Sequences scored: _____

Scoring of synovitis

	MCP-joints			
	2	**3**	**4**	**5**
Synovitis (0-3)				

Scoring of bone erosion and bone oedema

Bone erosion is scored 0-10, according to the proportion (in increments of 10%) of bone involved:
0: 0%, 1: 1-10%, 2: 11-20 %, ……., 10: 91-100%

Bone oedema is scored 0-3, according to the proportion (in increments of 33%) of bone involved:
0: 0%, 1: 1-33%, 2: 34-66 %, 3: 67-100%

Score from the articular surface (or its best estimated position if absent) to a depth of 1 cm.

		MCP joints			
		2	**3**	**4**	**5**
Bone erosion 0-10	**Proximal**				
	Distal				
Bone oedema 0-3	**Proximal**				
	Distal				

Appendix C

MRI Tenosynovitis Scoring

Patient Number: _____ ER / AJG

Wrist Flexor

	FCR	FPL	FDS/FDP	FCU
Score (0-3)				

Wrist Extensor

	I	II	III	IV	V	VI
Score (0-3)						

Hand flexor

	Thumb	Index	Middle	Ring	Little
Score (0-3)					

Score 0-3

Grade 0:	Normal
Grade 1:	<2mm thickening
Grade 2:	≥2mm and <5mm
Grade 3:	≥5mm thickening

List of References

1. Quinn, M.A. and P. Emery, *Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention.* Clin Exp Rheumatol, 2003. **21**(5 Suppl 31): p. S154-7.
2. Nell, V.P., et al., *Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis.* Rheumatology, 2004. **43**(7): p. 906-14.
3. van der Linden, M.P., et al., *Long-term impact of delay in assessment of patients with early arthritis.* Arthritis and rheumatism, 2010. **62**(12): p. 3537-46.
4. Lard, L.R., et al., *Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies.* Am J Med, 2001. **111**(6): p. 446-51.
5. Raza, K., et al., *Timing the therapeutic window of opportunity in early rheumatoid arthritis: proposal for definitions of disease duration in clinical trials.* Ann Rheum Dis, 2012. **71**(12): p. 1921-3.
6. Board, J., *Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3).* Heart, 2014. **100**(Suppl 2): p. ii1-ii67.
7. Hunt, L. and P. Emery, *Defining populations at risk of rheumatoid arthritis: the first steps to prevention.* Nat Rev Rheumatol, 2014.
8. Scott, D.L., et al., *The links between joint damage and disability in rheumatoid arthritis.* Rheumatology (Oxford), 2000. **39**(2): p. 122-32.
9. Bruynesteyn, K., et al., *Radiography as primary outcome in rheumatoid arthritis: acceptable sample sizes for trials with 3 months' follow up.* Ann Rheum Dis, 2004. **63**(11): p. 1413-8.
10. Sattar, N., et al., *Explaining how "high-grade" systemic inflammation accelerates vascular risk in rheumatoid arthritis.* Circulation, 2003. **108**(24): p. 2957-63.
11. Lindhardsen, J., et al., *The risk of myocardial infarction in rheumatoid arthritis and diabetes mellitus: a Danish nationwide cohort study.* Ann Rheum Dis, 2011. **70**(6): p. 929-34.
12. Markenson, J.A., *Worldwide trends in the socioeconomic impact and long-term prognosis of rheumatoid arthritis.* Semin Arthritis Rheum, 1991. **21**(2 Suppl 1): p. 4-12.
13. Lee, D.M. and M.E. Weinblatt, *Rheumatoid arthritis.* Lancet, 2001. **358**(9285): p. 903-11.
14. Finckh, A., H.K. Choi, and F. Wolfe, *Progression of radiographic joint damage in different eras: trends towards milder disease in rheumatoid arthritis are attributable to improved treatment.* Ann Rheum Dis, 2006. **65**(9): p. 1192-7.
15. Smolen, J.S., et al., *Treating rheumatoid arthritis to target: recommendations of an international task force.* Ann Rheum Dis, 2010. **69**(4): p. 631-7.
16. Silman, A.J. and J.E. Pearson, *Epidemiology and genetics of rheumatoid arthritis.* Arthritis Res, 2002. **4 Suppl 3**: p. S265-72.
17. Alamanos, Y., P.V. Voulgari, and A.A. Drosos, *Incidence and prevalence of rheumatoid arthritis, based on the 1987 American College of Rheumatology criteria: a systematic review.* Semin Arthritis Rheum, 2006. **36**(3): p. 182-8.

18. Del Puente, A., et al., *High incidence and prevalence of rheumatoid arthritis in Pima Indians.* Am J Epidemiol, 1989. **129**(6): p. 1170-8.
19. Harvey, J., et al., *Rheumatoid arthritis in a Chippewa Band. I. Pilot screening study of disease prevalence.* Arthritis Rheum, 1981. **24**(5): p. 717-21.
20. Symmons, D., et al., *The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century.* Rheumatology, 2002. **41**(7): p. 793-800.
21. Lawrence, J.S., *Prevalence of rheumatoid arthritis.* Ann Rheum Dis, 1961. **20**: p. 11-7.
22. Helmick, C.G., et al., *Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I.* Arthritis Rheum, 2008. **58**(1): p. 15-25.
23. McInnes, I.B. and G. Schett, *The pathogenesis of rheumatoid arthritis.* N Engl J Med, 2011. **365**(23): p. 2205-19.
24. Mankia, K. and P. Emery, *Is localized autoimmunity the trigger for rheumatoid arthritis? Unravelling new targets for prevention.* Discov Med, 2015. **20**(109): p. 129-35.
25. Franklin, E.C., et al., *An unusual protein component of high molecular weight in the serum of certain patients with rheumatoid arthritis.* J Exp Med, 1957. **105**(5): p. 425-38.
26. Waaler, E., *On the occurrence of a factor in human serum activating the specific agglutintion of sheep blood corpuscles. 1939.* APMIS, 2007. **115**(5): p. 422-38; discussion 439.
27. Wolfe, F., et al., *The prognosis of rheumatoid arthritis and undifferentiated polyarthritis syndrome in the clinic: a study of 1141 patients.* J Rheumatol, 1993. **20**(12): p. 2005-9.
28. Bukhari, M., et al., *Rheumatoid factor is the major predictor of increasing severity of radiographic erosions in rheumatoid arthritis: results from the Norfolk Arthritis Register Study, a large inception cohort.* Arthritis Rheum, 2002. **46**(4): p. 906-12.
29. Harrison, B. and D. Symmons, *Early inflammatory polyarthritis: results from the Norfolk Arthritis Register with a review of the literature. II. Outcome at three years.* Rheumatology (Oxford), 2000. **39**(9): p. 939-49.
30. Nell, V.P., et al., *Autoantibody profiling as early diagnostic and prognostic tool for rheumatoid arthritis.* Ann Rheum Dis, 2005. **64**(12): p. 1731-6.
31. Nienhuis, R.L. and E. Mandema, *A NEW SERUM FACTOR IN PATIENTS WITH RHEUMATOID ARTHRITIS; THE ANTIPERINUCLEAR FACTOR.* Ann Rheum Dis, 1964. **23**: p. 302-5.
32. Young, B.J., et al., *Anti-keratin antibodies in rheumatoid arthritis.* Br Med J, 1979. **2**(6182): p. 97-9.
33. Schellekens, G.A., et al., *Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies.* J Clin Invest, 1998. **101**(1): p. 273-81.
34. Rantapaa-Dahlqvist, S., et al., *Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis.* Arthritis Rheum, 2003. **48**(10): p. 2741-9.
35. van de Stadt, L.A., et al., *Development of the anti-citrullinated protein antibody repertoire prior to the onset of rheumatoid arthritis.* Arthritis and Rheumatism, 2011. **63**(11): p. 3226-33.

36. van der Woude, D., et al., *Epitope spreading of the anti-citrullinated protein antibody response occurs before disease onset and is associated with the disease course of early arthritis.* Ann Rheum Dis, 2010. **69**(8): p. 1554-61.
37. Ioan-Facsinay, A., et al., *Marked differences in fine specificity and isotype usage of the anti-citrullinated protein antibody in health and disease.* Arthritis and Rheumatism, 2008. **58**(10): p. 3000-3008.
38. Edwards, J.C. and G. Cambridge, *Sustained improvement in rheumatoid arthritis following a protocol designed to deplete B lymphocytes.* Rheumatology (Oxford), 2001. **40**(2): p. 205-11.
39. Kuhn, K.A., et al., *Antibodies against citrullinated proteins enhance tissue injury in experimental autoimmune arthritis.* J Clin Invest, 2006. **116**(4): p. 961-73.
40. Vossenaar, E.R., et al., *Citrullination of synovial proteins in murine models of rheumatoid arthritis.* Arthritis Rheum, 2003. **48**(9): p. 2489-500.
41. Hill, J.A., et al., *Arthritis induced by posttranslationally modified (citrullinated) fibrinogen in DR4-IE transgenic mice.* J Exp Med, 2008. **205**(4): p. 967-79.
42. Szanto, S., et al., *Induction of arthritis in HLA-DR4-humanized and HLA-DQ8-humanized mice by human cartilage proteoglycan aggrecan but only in the presence of an appropriate (non-MHC) genetic background.* Arthritis Rheum, 2004. **50**(6): p. 1984-95.
43. Hill, J.A., et al., *Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule.* J Immunol, 2003. **171**(2): p. 538-41.
44. Cantaert, T., et al., *Presence and role of anti-citrullinated protein antibodies in experimental arthritis models.* Arthritis Rheum, 2013. **65**(4): p. 939-48.
45. Trouw, L.A., et al., *Anti-cyclic citrullinated peptide antibodies from rheumatoid arthritis patients activate complement via both the classical and alternative pathways.* Arthritis Rheum, 2009. **60**(7): p. 1923-31.
46. Clavel, C., et al., *Induction of macrophage secretion of tumor necrosis factor alpha through Fcgamma receptor IIa engagement by rheumatoid arthritis-specific autoantibodies to citrullinated proteins complexed with fibrinogen.* Arthritis Rheum, 2008. **58**(3): p. 678-88.
47. Harre, U., et al., *Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin.* J Clin Invest, 2012. **122**(5): p. 1791-802.
48. Wigerblad, G., et al., *Autoantibodies to citrullinated proteins induce joint pain independent of inflammation via a chemokine-dependent mechanism.* 2016. **75**(4): p. 730-8.
49. Catrina, A.I., et al., *Mechanisms leading from systemic autoimmunity to joint-specific disease in rheumatoid arthritis.* Nat Rev Rheumatol, 2017. **13**(2): p. 79-86.
50. MacGregor, A.J., et al., *Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins.* Arthritis Rheum, 2000. **43**(1): p. 30-7.
51. Lin, J.P., et al., *Familial clustering of rheumatoid arthritis with other autoimmune diseases.* Hum Genet, 1998. **103**(4): p. 475-82.

52. Frisell, T., et al., *Familial risks and heritability of rheumatoid arthritis: role of rheumatoid factor/anti-citrullinated protein antibody status, number and type of affected relatives, sex, and age.* Arthritis Rheum, 2013. **65**(11): p. 2773-82.
53. Speed, D., et al., *Improved heritability estimation from genome-wide SNPs.* Am J Hum Genet, 2012. **91**(6): p. 1011-21.
54. Stahl, E.A., et al., *Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci.* Nat Genet, 2010. **42**(6): p. 508-14.
55. Eyre, S., et al., *High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis.* Nat Genet, 2012. **44**(12): p. 1336-40.
56. Okada, Y., et al., *Genetics of rheumatoid arthritis contributes to biology and drug discovery.* Nature, 2014. **506**(7488): p. 376-81.
57. Balandraud, N., et al., *HLA-DRB1 genotypes and the risk of developing anti citrullinated protein antibody (ACPA) positive rheumatoid arthritis.* PLoS One, 2013. **8**(5): p. e64108.
58. Raychaudhuri, S., et al., *Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis.* Nat Genet, 2012. **44**(3): p. 291-6.
59. Gregersen, P.K., J. Silver, and R.J. Winchester, *The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis.* Arthritis Rheum, 1987. **30**(11): p. 1205-13.
60. De Almeida, D.E., et al., *Immune dysregulation by the rheumatoid arthritis shared epitope.* J Immunol, 2010. **185**(3): p. 1927-34.
61. Huizinga, T.W., et al., *Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins.* Arthritis Rheum, 2005. **52**(11): p. 3433-8.
62. Harrison, B., et al., *The influence of HLA-DRB1 alleles and rheumatoid factor on disease outcome in an inception cohort of patients with early inflammatory arthritis.* Arthritis Rheum, 1999. **42**(10): p. 2174-83.
63. Hitchon, C.A., et al., *Antibodies to porphyromonas gingivalis are associated with anticitrullinated protein antibodies in patients with rheumatoid arthritis and their relatives.* J Rheumatol, 2010. **37**(6): p. 1105-12.
64. Begovich, A.B., et al., *A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis.* Am J Hum Genet, 2004. **75**(2): p. 330-7.
65. Hinks, A., et al., *Association between the PTPN22 gene and rheumatoid arthritis and juvenile idiopathic arthritis in a UK population: further support that PTPN22 is an autoimmunity gene.* Arthritis Rheum, 2005. **52**(6): p. 1694-9.
66. Plenge, R.M., et al., *Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4.* Am J Hum Genet, 2005. **77**(6): p. 1044-60.
67. Criswell, L.A., et al., *Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes.* Am J Hum Genet, 2005. **76**(4): p. 561-71.

68. Eyre, S., et al., *Overlapping genetic susceptibility variants between three autoimmune disorders: rheumatoid arthritis, type 1 diabetes and coeliac disease.* Arthritis Res Ther, 2010. **12**(5): p. R175.
69. Verpoort, K.N., et al., *Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis.* Arthritis Rheum, 2005. **52**(10): p. 3058-62.
70. Sigurdsson, S., et al., *Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis.* Arthritis Rheum, 2007. **56**(7): p. 2202-10.
71. Padyukov, L., et al., *A genome-wide association study suggests contrasting associations in ACPA-positive versus ACPA-negative rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2010.
72. Lichtenstein, P., et al., *The Swedish Twin Registry in the third millennium: an update.* Twin Res Hum Genet, 2006. **9**(6): p. 875-82.
73. Magnusson, P.K., et al., *The Swedish Twin Registry: establishment of a biobank and other recent developments.* Twin Res Hum Genet, 2013. **16**(1): p. 317-29.
74. Hensvold, A.H., et al., *Environmental and genetic factors in the development of anticitrullinated protein antibodies (ACPAs) and ACPA-positive rheumatoid arthritis: an epidemiological investigation in twins.* Ann Rheum Dis, 2015. **74**(2): p. 375-80.
75. Jaenisch, R. and A. Bird, *Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals.* Nat Genet, 2003. **33 Suppl**: p. 245-54.
76. Ospelt, C., et al., *Inflammatory memories: is epigenetics the missing link to persistent stromal cell activation in rheumatoid arthritis?* Autoimmun Rev, 2011. **10**(9): p. 519-24.
77. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-5.
78. Viatte, S., D. Plant, and S. Raychaudhuri, *Genetics and epigenetics of rheumatoid arthritis.* Nat Rev Rheumatol, 2013. **9**(3): p. 141-53.
79. Horiuchi, M., et al., *Expression and function of histone deacetylases in rheumatoid arthritis synovial fibroblasts.* J Rheumatol, 2009. **36**(8): p. 1580-9.
80. Nakano, K., et al., *DNA methylome signature in rheumatoid arthritis.* Ann Rheum Dis, 2013. **72**(1): p. 110-7.
81. Karouzakis, E., et al., *DNA hypomethylation in rheumatoid arthritis synovial fibroblasts.* Arthritis Rheum, 2009. **60**(12): p. 3613-22.
82. Kim, V.N., *MicroRNA biogenesis: coordinated cropping and dicing.* Nat Rev Mol Cell Biol, 2005. **6**(5): p. 376-85.
83. Bartel, D.P., *MicroRNAs: genomics, biogenesis, mechanism, and function.* Cell, 2004. **116**(2): p. 281-97.
84. Dennis, C., *Small RNAs: The genome's guiding hand?* Nature, 2002. **420**(6917): p. 732-732.
85. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.* Cell, 2005. **120**(1): p. 15-20.
86. Churov, A.V., E.K. Oleinik, and M. Knip, *MicroRNAs in rheumatoid arthritis: altered expression and diagnostic potential.* Autoimmun Rev, 2015. **14**(11): p. 1029-37.

87. Stanczyk, J., et al., *Altered expression of microRNA-203 in rheumatoid arthritis synovial fibroblasts and its role in fibroblast activation.* Arthritis Rheum, 2011. **63**(2): p. 373-81.
88. Stanczyk, J., et al., *Altered expression of MicroRNA in synovial fibroblasts and synovial tissue in rheumatoid arthritis.* Arthritis & Rheumatism, 2008. **58**(4): p. 1001-1009.
89. Nakasa, T., et al., *Expression of microRNA-146 in rheumatoid arthritis synovial tissue.* Arthritis Rheum, 2008. **58**(5): p. 1284-92.
90. Miller, F.W., et al., *Epidemiology of environmental exposures and human autoimmune diseases: findings from a National Institute of Environmental Health Sciences Expert Panel Workshop.* J Autoimmun, 2012. **39**(4): p. 259-71.
91. Lahiri, M., et al., *Using lifestyle factors to identify individuals at higher risk of inflammatory polyarthritis (results from the European Prospective Investigation of Cancer-Norfolk and the Norfolk Arthritis Register--the EPIC-2-NOAR Study).* Ann Rheum Dis, 2014. **73**(1): p. 219-26.
92. Crowson, C.S., et al., *Contribution of obesity to the rise in incidence of rheumatoid arthritis.* Arthritis Care and Research, 2013. **65**(1): p. 71-77.
93. Heliovaara, M., et al., *Coffee consumption, rheumatoid factor, and the risk of rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2000. **59**(8): p. 631-635.
94. Karlson, E.W., et al., *Coffee Consumption and Risk of Rheumatoid Arthritis.* Arthritis and Rheumatism, 2003. **48**(11): p. 3055-3060.
95. Jacobsson, L.T.H., et al., *Perinatal characteristics and risk of rheumatoid arthritis.* British Medical Journal, 2003. **326**(7398): p. 1068-1069.
96. Doran, M.F., et al., *The Effect of Oral Contraceptives and Estrogen Replacement Therapy on the Risk of Rheumatoid Arthritis: A Population Based Study.* Journal of Rheumatology, 2004. **31**(2): p. 207-213.
97. Karlson, E.W., et al., *Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? Results from the nurses' health study.* Arthritis and Rheumatism, 2004. **50**(11): p. 3458-3467.
98. Pattison, D.J., R.A. Harrison, and D.P.M. Symmons, *The role of diet in susceptibility to rheumatoid arthritis: a systematic review.* Journal of Rheumatology, 2004. **31**(7): p. 1310-9.
99. Stolt, P., et al., *Silica exposure is associated with increased risk of developing rheumatoid arthritis: Results from the Swedish EIRA study.* Annals of the Rheumatic Diseases, 2005. **64**(4): p. 582-586.
100. Bengtsson, C., et al., *Socioeconomic status and the risk of developing rheumatoid arthritis: Results from the Swedish EIRA study.* Annals of the Rheumatic Diseases, 2005. **64**(11): p. 1588-1594.
101. Benito-Garcia, E., et al., *Protein, iron, and meat consumption and risk for rheumatoid arthritis: a prospective cohort study.* Arthritis research & therapy, 2007. **9**(1): p. R16.
102. Costenbader, K.H., et al., *Vitamin D intake and risks of systemic lupus erythematosus and rheumatoid arthritis in women.* Annals of the Rheumatic Diseases, 2008. **67**(4): p. 530-535.
103. Mandl, L.A., et al., *Is birthweight associated with risk of rheumatoid arthritis? data from a large cohort Study.* Annals of the Rheumatic Diseases, 2009. **68**(4): p. 514-518.
104. Bengtsson, C., et al., *Psychosocial stress at work and the risk of developing rheumatoid arthritis: Results from the Swedish EIRA study.* Psychotherapy and Psychosomatics, 2009. **78**(3): p. 193-194.

105. Pikwer, M., et al., *Breast feeding, but not use of oral contraceptives, is associated with a reduced risk of rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2009. **68**(4): p. 526-530.
106. Rosell, M., et al., *Dietary fish and fish oil and the risk of rheumatoid arthritis.* Epidemiology, 2009. **20**(6): p. 896-901.
107. Stolt, P., et al., *Silica exposure among male current smokers is associated with a high risk of developing ACPA-positive rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2010. **69**(6): p. 1072-1076.
108. Arkema, E.V., E.W. Karlson, and K.H. Costenbader, *A prospective study of periodontal disease and risk of rheumatoid arthritis.* Journal of Rheumatology, 2010. **37**(9): p. 1800-1804.
109. Lahiri, M., et al., *Higher fruit and fructose consumption is associated with a reduced risk of inflammatory polyarthritis and rheumatoid arthritis in men. Results from the European prospective investigation of cancer (Norfolk) and the norfolk arthritis register.* Arthritis and Rheumatism, 2011. **1)**.
110. Parks, C.G., et al., *Childhood socioeconomic factors and perinatal characteristics influence development of rheumatoid arthritis in adulthood.* Annals of the Rheumatic Diseases, 2013. **72**(3): p. 350-6.
111. Arkema, E.V., et al., *Exposure to ultraviolet-B and risk of developing rheumatoid arthritis among women in the Nurses' Health Study.* Annals of the Rheumatic Diseases, 2013. **72**(4): p. 506-511.
112. Scott, I.C., et al., *The protective effect of alcohol on developing rheumatoid arthritis: a systematic review and meta-analysis.* Rheumatology, 2013. **52**(5): p. 856-67.
113. Oliver, J.E. and A.J. Silman, *Why are women predisposed to autoimmune rheumatic diseases?* Arthritis Res Ther, 2009. **11**(5): p. 252.
114. Jorgensen, C., et al., *Oral contraception, parity, breast feeding, and severity of rheumatoid arthritis.* Ann Rheum Dis, 1996. **55**(2): p. 94-8.
115. Pedersen, M., et al., *Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides.* Arthritis Res Ther, 2006. **8**(4): p. R133.
116. Ostensen, M. and P.M. Villiger, *The remission of rheumatoid arthritis during pregnancy.* Semin Immunopathol, 2007. **29**(2): p. 185-91.
117. de Man, Y.A., et al., *Disease activity of rheumatoid arthritis during pregnancy: results from a nationwide prospective study.* Arthritis Rheum, 2008. **59**(9): p. 1241-8.
118. Sugiyama, D., et al., *Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies.* Annals of the Rheumatic Diseases, 2010. **69**(1): p. 70-81.
119. Klareskog, L., et al., *A new model for an etiology of rheumatoid arthritis: Smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination.* Arthritis and Rheumatism, 2006. **54**(1): p. 38-46.
120. Kallberg, H., et al., *Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis.* Am J Hum Genet, 2007. **80**(5): p. 867-75.
121. Demoruelle, M.K., et al., *Brief report: airways abnormalities and rheumatoid arthritis-related autoantibodies in subjects without arthritis: early injury or initiating site of autoimmunity?* Arthritis and rheumatism, 2012. **64**(6): p. 1756-1761.

122. Reynisdottir, G., et al., *Structural changes and antibody enrichment in the lungs are early features of anti-citrullinated protein antibody-positive rheumatoid arthritis.* Arthritis Rheumatol, 2014. **66**(1): p. 31-9.
123. Rangel-Moreno, J., et al., *Inducible bronchus-associated lymphoid tissue (iBALT) in patients with pulmonary complications of rheumatoid arthritis.* J Clin Invest, 2006. **116**(12): p. 3183-94.
124. Makrygiannakis, D., et al., *Smoking increases peptidylarginine deiminase 2 enzyme expression in human lungs and increases citrullination in BAL cells.* Ann Rheum Dis, 2008. **67**(10): p. 1488-92.
125. Ytterberg, A.J., et al., *Shared immunological targets in the lungs and joints of patients with rheumatoid arthritis: identification and validation.* Ann Rheum Dis, 2015. **74**(9): p. 1772-7.
126. Reynisdottir, G., et al., *Signs of immune activation and local inflammation are present in the bronchial tissue of patients with untreated early rheumatoid arthritis.* Ann Rheum Dis, 2016. **75**(9): p. 1722-7.
127. Willis, V.C., et al., *Sputa autoantibodies in patients with established rheumatoid arthritis and subjects at-risk for future clinically apparent disease.* Arthritis Rheum, 2013.
128. Rosenstein, E.D., et al., *Hypothesis: the humoral immune response to oral bacteria provides a stimulus for the development of rheumatoid arthritis.* Inflammation, 2004. **28**(6): p. 311-8.
129. Wegner, N., et al., *Peptidylarginine deiminase from Porphyromonas gingivalis citrullinates human fibrinogen and alpha-enolase: implications for autoimmunity in rheumatoid arthritis.* Arthritis and rheumatism, 2010. **62**(9): p. 2662-72.
130. de Pablo, P., T. Dietrich, and T.E. McAlindon, *Association of periodontal disease and tooth loss with rheumatoid arthritis in the US population.* The Journal of rheumatology, 2008. **35**(1): p. 70-6.
131. Demmer, R.T., et al., *Periodontal disease, tooth loss and incident rheumatoid arthritis: results from the First National Health and Nutrition Examination Survey and its epidemiological follow-up study.* Journal of Clinical Periodontology, 2011. **38**(11): p. 998-1006.
132. Chen, H.H., et al., *Association between a history of periodontitis and the risk of rheumatoid arthritis: a nationwide, population-based, case-control study.* Annals of the Rheumatic Diseases, 2012. **31**: p. 31.
133. Hitchon, C.A., et al., *Antibodies to Porphyromonas gingivalis are associated with anticitrullinated protein antibodies in patients with rheumatoid arthritis and their relatives.* Journal of Rheumatology, 2010. **37**(6): p. 1105-1112.
134. De Smit, M.J., et al., *Antibodies against porphyromonas gingivalis correlate with rheumatoid arthritis-specific auto-immunity in arthralgia patients.* Arthritis and Rheumatism, 2012. **64**: p. S514.
135. Mikuls, T.R., et al., *Porphyromonas gingivalis and disease-related autoantibodies in individuals at increased risk of rheumatoid arthritis.* Arthritis and Rheumatism, 2012. **64**(11): p. 3522-3530.
136. Round, J.L. and S.K. Mazmanian, *The gut microbiota shapes intestinal immune responses during health and disease.* Nat Rev Immunol, 2009. **9**(5): p. 313-23.
137. Scher, J.U. and S.B. Abramson, *The microbiome and rheumatoid arthritis.* Nature reviews. Rheumatology, 2011. **7**(10): p. 569-78.

138. Newkirk, M.M., et al., *Distinct bacterial colonization patterns of Escherichia coli subtypes associate with rheumatoid factor status in early inflammatory arthritis.* Rheumatology (Oxford), 2010. **49**(7): p. 1311-6.
139. Ball, R.J., et al., *Systematic review and meta-analysis of the sero-epidemiological association between Epstein-Barr virus and rheumatoid arthritis.* Arthritis Res Ther, 2015. **17**: p. 274.
140. Wilson, C., H. Tiwana, and A. Ebringer, *Molecular mimicry between HLA-DR alleles associated with rheumatoid arthritis and Proteus mirabilis as the Aetiological basis for autoimmunity.* Microbes Infect, 2000. **2**(12): p. 1489-96.
141. Newkirk, M.M., et al., *Elevated levels of IgM and IgA antibodies to Proteus mirabilis and IgM antibodies to Escherichia coli are associated with early rheumatoid factor (RF)-positive rheumatoid arthritis.* Rheumatology (Oxford), 2005. **44**(11): p. 1433-41.
142. Auger, I. and J. Roudier, *A function for the QKRAA amino acid motif: mediating binding of DnaJ to DnaK. Implications for the association of rheumatoid arthritis with HLA-DR4.* J Clin Invest, 1997. **99**(8): p. 1818-22.
143. Kohashi, O., et al., *Susceptibility to adjuvant-induced arthritis among germfree, specific-pathogen-free, and conventional rats.* Infect Immun, 1979. **26**(3): p. 791-4.
144. Kohashi, O., et al., *Reverse effect of gram-positive bacteria vs. gram-negative bacteria on adjuvant-induced arthritis in germfree rats.* Microbiol Immunol, 1985. **29**(6): p. 487-97.
145. Kohashi, O., et al., *Suppressive effect of Escherichia coli on adjuvant-induced arthritis in germ-free rats.* Arthritis Rheum, 1986. **29**(4): p. 547-53.
146. Sinkorova, Z., et al., *Commensal intestinal bacterial strains trigger ankylosing enthesopathy of the ankle in inbred B10.BR (H-2(k)) male mice.* Hum Immunol, 2008. **69**(12): p. 845-50.
147. Abdollahi-Roodsaz, S., et al., *Stimulation of TLR2 and TLR4 differentially skews the balance of T cells in a mouse model of arthritis.* J Clin Invest, 2008. **118**(1): p. 205-16.
148. Sandhya, P., et al., *Does the buck stop with the bugs?: an overview of microbial dysbiosis in rheumatoid arthritis.* Int J Rheum Dis, 2016. **19**(1): p. 8-20.
149. Zhang, X., et al., *The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment.* Nat Med, 2015. **21**(8): p. 895-905.
150. Scher, J.U., et al., *Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis.* Elife, 2013. **2**: p. e01202.
151. *A framework for human microbiome research.* Nature, 2012. **486**(7402): p. 215-21.
152. Qin, J., et al., *A human gut microbial gene catalogue established by metagenomic sequencing.* Nature, 2010. **464**(7285): p. 59-65.
153. Gerlag, D.M., et al., *EULAR recommendations for terminology and research in individuals at risk of rheumatoid arthritis: report from the Study Group for Risk Factors for Rheumatoid Arthritis.* Ann Rheum Dis, 2012. **71**(5): p. 638-41.
154. van Aken, J., et al., *Comparison of long term outcome of patients with rheumatoid arthritis presenting with undifferentiated arthritis or with*

rheumatoid arthritis: an observational cohort study. Ann Rheum Dis, 2006. **65**(1): p. 20-5.
155. McNally, E., et al., *Diagnostic accuracy of a clinical prediction rule (CPR) for identifying patients with recent-onset undifferentiated arthritis who are at a high risk of developing rheumatoid arthritis: a systematic review and meta-analysis.* Semin Arthritis Rheum, 2014. **43**(4): p. 498-507.
156. Van Der Helm-Van Mil, A.H.M., et al., *A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: How to guide individual treatment decisions.* Arthritis and Rheumatism, 2007. **56**(2): p. 433-440.
157. Jansen, L.M.A., et al., *One year outcome of undifferentiated polyarthritis.* Annals of the Rheumatic Diseases, 2002. **61**(8): p. 700-703.
158. Aletaha, D., et al., *2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative.* Arthritis and rheumatism, 2010. **62**(9): p. 2569-81.
159. Arnett, F.C., et al., *The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis.* Arthritis Rheum, 1988. **31**(3): p. 315-24.
160. Aletaha, D., F.C. Breedveld, and J.S. Smolen, *The need for new classification criteria for rheumatoid arthritis.* Arthritis Rheum, 2005. **52**(11): p. 3333-6.
161. Radner, H., et al., *Performance of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: a systematic literature review.* Ann Rheum Dis, 2014. **73**(1): p. 114-23.
162. van Gaalen, F., et al., *The devil in the details: the emerging role of anticitrulline autoimmunity in rheumatoid arthritis.* J Immunol, 2005. **175**(9): p. 5575-80.
163. Humphreys, J.H., et al., *The incidence of rheumatoid arthritis in the UK: comparisons using the 2010 ACR/EULAR classification criteria and the 1987 ACR classification criteria. Results from the Norfolk Arthritis Register.* Ann Rheum Dis, 2013. **72**(8): p. 1315-20.
164. Hunt, L. and M. Buch, *The 'therapeutic window' and treating to target in rheumatoid arthritis.* Clin Med, 2013. **13**(4): p. 387-90.
165. 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.
166. Smolen, J.S., D. Aletaha, and I.B. McInnes, *Rheumatoid arthritis.* Lancet, 2016. **388**(10055): p. 2023-2038.
167. Wolfe, F. and D.J. Hawley, *Remission in rheumatoid arthritis.* J Rheumatol, 1985. **12**(2): p. 245-52.
168. van der Heijde, D.M., et al., *Influence of prognostic features on the final outcome in rheumatoid arthritis: a review of the literature.* Semin Arthritis Rheum, 1988. **17**(4): p. 284-92.
169. Weyand, C.M., et al., *The influence of sex on the phenotype of rheumatoid arthritis.* Arthritis Rheum, 1998. **41**(5): p. 817-22.
170. Ahlmen, M., et al., *Influence of gender on assessments of disease activity and function in early rheumatoid arthritis in relation to radiographic joint damage.* Ann Rheum Dis, 2010. **69**(1): p. 230-3.
171. Masdottir, B., et al., *Smoking, rheumatoid factor isotypes and severity of rheumatoid arthritis.* Rheumatology (Oxford), 2000. **39**(11): p. 1202-5.

172. Wolfe, F., *The effect of smoking on clinical, laboratory, and radiographic status in rheumatoid arthritis.* J Rheumatol, 2000. **27**(3): p. 630-7.
173. 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.
174. van Leeuwen, M.A., et al., *Individual relationship between progression of radiological damage and the acute phase response in early rheumatoid arthritis. Towards development of a decision support system.* J Rheumatol, 1997. **24**(1): p. 20-7.
175. Eberhardt, K. and E. Fex, *Clinical course and remission rate in patients with early rheumatoid arthritis: relationship to outcome after 5 years.* Br J Rheumatol, 1998. **37**(12): p. 1324-9.
176. Jayakumar, K., et al., *Sustained clinical remission in rheumatoid arthritis: prevalence and prognostic factors in an inception cohort of patients treated with conventional DMARDS.* Rheumatology (Oxford), 2012. **51**(1): p. 169-75.
177. Gerlag, D.M., J.M. Norris, and P.P. Tak, *Towards prevention of autoantibody-positive rheumatoid arthritis: from lifestyle modification to preventive treatment.* Rheumatology (Oxford), 2016. **55**(4): p. 607-14.
178. Deane, K.D., *Can rheumatoid arthritis be prevented?* Best Pract Res Clin Rheumatol, 2013. **27**(4): p. 467-85.
179. del Puente, A., et al., *The incidence of rheumatoid arthritis is predicted by rheumatoid factor titer in a longitudinal population study.* Arthritis Rheum, 1988. **31**(10): p. 1239-44.
180. van Steenbergen, H.W., T.W. Huizinga, and A.H. van der Helm-van Mil, *Review: The Preclinical Phase of Rheumatoid Arthritis: What Is Acknowledged and What Needs to be Assessed?* Arthritis Rheum, 2013. **65**(9): p. 2219-32.
181. Aho, K., et al., *Antikeratin antibody and antiperinuclear factor as markers for subclinical rheumatoid disease process.* Journal of Rheumatology, 1993. **20**(8): p. 1278-81.
182. Aho, K., et al., *Antifilaggrin antibodies within 'normal' range predict rheumatoid arthritis in a linear fashion.* Journal of Rheumatology, 2000. **27**(12): p. 2743-2746.
183. Chibnik, L.B., et al., *Comparison of threshold cutpoints and continuous measures of anti-cyclic citrullinated peptide antibodies in predicting future rheumatoid arthritis.* The Journal of rheumatology, 2009. **36**(4): p. 706-711.
184. Kokkonen, H., et al., *Antibodies of IgG, IgA and IgM isotypes against cyclic citrullinated peptide precede the development of rheumatoid arthritis.* Arthritis research & therapy, 2011. **13**(1): p. R13.
185. Majka, D.S., et al., *Duration of preclinical rheumatoid arthritis-related autoantibody positivity increases in subjects with older age at time of disease diagnosis.* Annals of the Rheumatic Diseases, 2008. **67**(6): p. 801-807.
186. Jonsson, T., et al., *Population study of the importance of rheumatoid factor isotypes in adults.* Annals of the Rheumatic Diseases, 1992. **51**(7): p. 863-868.
187. Nielen, M.M., et al., *Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors.* Arthritis Rheum, 2004. **50**(2): p. 380-6.

188. Sokolove, J., et al., *Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis.* PLoS ONE, 2012. **7**(5).
189. Brink, M., et al., *Multiplex analyses of antibodies against citrullinated peptides in individuals prior to development of rheumatoid arthritis.* Arthritis & Rheumatism, 2013. **65**(4): p. 899-910.
190. Karlson, E.W., et al., *Biomarkers of inflammation and development of rheumatoid arthritis in women from two prospective cohort studies.* Arthritis and Rheumatism, 2009. **60**(3): p. 641-652.
191. Jorgensen, K.T., et al., *Cytokines, autoantibodies and viral antibodies in premorbid and postdiagnostic sera from patients with rheumatoid arthritis: Case-control study nested in a cohort of Norwegian blood donors.* Annals of the Rheumatic Diseases, 2008. **67**(6): p. 860-866.
192. Aho, K., et al., *Serum immunoglobulins and the risk of rheumatoid arthritis.* Annals of the Rheumatic Diseases, 1997. **56**(6): p. 351-356.
193. Aho, K., et al., *Serum C-reactive protein does not predict rheumatoid arthritis.* Journal of Rheumatology, 2000. **27**(5): p. 1136-1138.
194. Nielen, M.M.J., et al., *Increased levels of C-reactive protein in serum from blood donors before the onset of rheumatoid arthritis.* Arthritis and Rheumatism, 2004. **50**(8): p. 2423-2427.
195. Deane, K.D., et al., *The number of elevated cytokines and chemokines in preclinical seropositive rheumatoid arthritis predicts time to diagnosis in an age-dependent manner.* Arthritis and Rheumatism, 2010. **62**(11): p. 3161-3172.
196. Kokkonen, H., et al., *Up-regulation of cytokines and chemokines predates the onset of rheumatoid arthritis.* Arthritis and Rheumatism, 2010. **62**(2): p. 383-391.
197. Rantapaa-Dahlqvist, S., et al., *Up regulation of monocyte chemoattractant protein-1 expression in anti-citrulline antibody and immunoglobulin M rheumatoid factor positive subjects precedes onset of inflammatory response and development of overt rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2007. **66**(1): p. 121-123.
198. Nielen, M.M.J., et al., *Simultaneous development of acute phase response and autoantibodies in preclinical rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2006. **65**(4): p. 535-537.
199. van Schaardenburg, D., et al., *Bone metabolism is altered in preclinical rheumatoid arthritis.* Ann Rheum Dis, 2011. **70**(6): p. 1173-4.
200. Hiraki, L.T., et al., *Circulating 25-hydroxyvitamin D level and risk of developing rheumatoid arthritis.* Rheumatology (Oxford), 2014. **53**(12): p. 2243-8.
201. Silman, A.J., E. Hennessy, and B. Ollier, *Incidence of rheumatoid arthritis in a genetically predisposed population.* Br J Rheumatol, 1992. **31**(6): p. 365-8.
202. El-Gabalawy, H.S., et al., *Immunogenetic risks of anti-cyclical citrullinated peptide antibodies in a North American Native population with rheumatoid arthritis and their first-degree relatives.* Journal of Rheumatology, 2009. **36**(6): p. 1130-1135.
203. Kolfenbach, J.R., et al., *A prospective approach to investigating the natural history of preclinical rheumatoid arthritis (RA) using first-degree relatives of probands with RA.* Arthritis Care and Research, 2009. **61**(12): p. 1735-1742.

204. Arlestig, L., et al., *Antibodies against cyclic citrullinated peptides of IgG, IgA and IgM isotype and rheumatoid factor of IgM and IgA isotype are increased in unaffected members of multicase rheumatoid arthritis families from northern Sweden.* Annals of the Rheumatic Diseases, 2012. **71**(6): p. 825-829.
205. Barra, L., et al., *Anti-citrullinated protein antibodies in unaffected first-degree relatives of rheumatoid arthritis patients.* Arthritis and Rheumatism, 2013. **65**(6): p. 1439-1447.
206. Demoruelle, M.K., et al., *Performance of anti-cyclic citrullinated peptide assays differs in subjects at increased risk of rheumatoid arthritis and subjects with established disease.* Arthritis and Rheumatism, 2013. **65**(9): p. 2243-2252.
207. Young, K.A., et al., *Relatives without rheumatoid arthritis show reactivity to anti-citrullinated protein/peptide antibodies that are associated with arthritis-related traits: Studies of the etiology of rheumatoid arthritis.* Arthritis and Rheumatism, 2013. **65**(8): p. 1995-2004.
208. Nielen, M.M.J., et al., *Specific Autoantibodies Precede the Symptoms of Rheumatoid Arthritis: A Study of Serial Measurements in Blood Donors.* Arthritis and Rheumatism, 2004. **50**(2): p. 380-386.
209. El-Gabalawy, H.S., et al., *Familial clustering of the serum cytokine profile in the relatives of rheumatoid arthritis patients.* Arthritis and Rheumatism, 2012. **64**(6): p. 1720-1729.
210. Hughes-Austin, J.M., et al., *Multiple cytokines and chemokines are associated with rheumatoid arthritis-related autoimmunity in first-degree relatives without rheumatoid arthritis: Studies of the Aetiology of Rheumatoid Arthritis (SERA).* Annals of the Rheumatic Diseases, 2013. **72**(6): p. 901-907.
211. Gan, R.W., et al., *Lower omega-3 fatty acids are associated with the presence of anti-cyclic citrullinated peptide autoantibodies in a population at risk for future rheumatoid arthritis: a nested case-control study.* Rheumatology (Oxford), 2016. **55**(2): p. 367-76.
212. Research, N.I.f.H. *PREVeNT RA - PRe-clinical EValuation of Novel Targets in RA*. 24.02.2017]; Available from: https://www.aruk.manchester.ac.uk/tacera_preventra/.
213. Knip, M., et al., *Dietary intervention in infancy and later signs of beta-cell autoimmunity.* N Engl J Med, 2010. **363**(20): p. 1900-8.
214. Schmid, S., et al., *BABYDIET, a feasibility study to prevent the appearance of islet autoantibodies in relatives of patients with Type 1 diabetes by delaying exposure to gluten.* Diabetologia, 2004. **47**(6): p. 1130-1.
215. Gale, E.A., et al., *European Nicotinamide Diabetes Intervention Trial (ENDIT): a randomised controlled trial of intervention before the onset of type 1 diabetes.* Lancet, 2004. **363**(9413): p. 925-31.
216. Bos, W.H., et al., *Arthritis development in patients with arthralgia is strongly associated with anti-citrullinated protein antibody status: A prospective cohort study.* Annals of the Rheumatic Diseases, 2010. **69**(3): p. 490-494.
217. Rakieh, C., et al., *Predicting the development of clinical arthritis in anti-CCP positive individuals with non-specific musculoskeletal symptoms: a prospective observational cohort study.* Ann Rheum Dis, 2014.
218. *Biomarkers and surrogate endpoints: preferred definitions and conceptual framework.* Clin Pharmacol Ther, 2001. **69**(3): p. 89-95.

219. Illei, G.G., et al., *Biomarkers in systemic lupus erythematosus: I. General overview of biomarkers and their applicability.* Arthritis & Rheumatism, 2004. **50**(6): p. 1709-1720.
220. Van De Stadt, L.A., et al., *The extent of the anti-citrullinated protein antibody repertoire is associated with arthritis development in patients with seropositive arthralgia.* Annals of the Rheumatic Diseases, 2011. **70**(1): p. 128-133.
221. Shi, J., et al., *Anti-carbamylated protein antibodies are present in arthralgia patients and predict the development of rheumatoid arthritis.* Arthritis and Rheumatism, 2013. **65**(4): p. 911-915.
222. Rombouts, Y., et al., *Anti-citrullinated protein antibodies acquire a pro-inflammatory Fc glycosylation phenotype prior to the onset of rheumatoid arthritis.* Ann Rheum Dis, 2015. **74**(1): p. 234-41.
223. Van Baarsen, L.G.M., et al., *Gene expression profiling in autoantibody-positive patients with arthralgia predicts development of arthritis.* Arthritis and Rheumatism, 2010. **62**(3): p. 694-704.
224. Lubbers, J., et al., *The type I IFN signature as a biomarker of preclinical rheumatoid arthritis.* Ann Rheum Dis, 2013. **72**(5): p. 776-80.
225. Lubbers, J., et al., *B cell signature contributes to the prediction of RA development in patients with arthralgia.* Ann Rheum Dis, 2015. **74**(9): p. 1786-8.
226. Limper, M., et al., *The acute-phase response is not predictive for the development of arthritis in seropositive arthralgia - A prospective cohort study.* Journal of Rheumatology, 2012. **39**(10): p. 1914-1917.
227. Van De Stadt, L.A., et al., *Dyslipidaemia in patients with seropositive arthralgia predicts the development of arthritis.* Annals of the Rheumatic Diseases, 2012. **71**(11): p. 1915-1916.
228. Maijer, K.I., et al., *Serum Vaspin Levels Are Associated with the Development of Clinically Manifest Arthritis in Autoantibody-Positive Individuals.* PLoS One, 2015. **10**(12): p. e0144932.
229. Van De Sande, M.G.H., et al., *Different stages of rheumatoid arthritis: Features of the synovium in the preclinical phase.* Annals of the Rheumatic Diseases, 2011. **70**(5): p. 772-777.
230. de Hair, M.J., et al., *Expression of Prostaglandin E2 Enzymes in the Synovium of Arthralgia Patients at Risk of Developing Rheumatoid Arthritis and in Early Arthritis Patients.* PLoS One, 2015. **10**(7): p. e0133669.
231. Ramwadhdoebe, T.H., et al., *Human lymph-node CD8(+) T cells display an altered phenotype during systemic autoimmunity.* Clin Transl Immunology, 2016. **5**(4): p. e67.
232. Rodriguez-Carrio, J., et al., *Altered Innate Lymphoid Cells subsets in human lymph node biopsies during the at risk and earliest phase of rheumatoid arthritis.* Arthritis Rheumatol, 2016.
233. Krabben, A., et al., *MRI of hand and foot joints of patients with anticitrullinated peptide antibody positive arthralgia without clinical arthritis.* Ann Rheum Dis, 2013. **72**(9): p. 1540-4.
234. Gent, Y.Y., et al., *Three-year clinical outcome following baseline magnetic resonance imaging in anti-citrullinated protein antibody-positive arthralgia patients: an exploratory study.* Arthritis Rheumatol, 2014. **66**(10): p. 2909-10.

235. Kleyer, A., et al., *High prevalence of tenosynovial inflammation before onset of rheumatoid arthritis and its link to progression to RA-A combined MRI/CT study.* Semin Arthritis Rheum, 2016.
236. van de Stadt, L.A., et al., *The value of ultrasonography in predicting arthritis in auto-antibody positive arthralgia patients: a prospective cohort study.* Arthritis research & therapy, 2010. **12**(3): p. R98.
237. Nam, J.L., et al., *Ultrasound findings predict progression to inflammatory arthritis in anti-CCP antibody-positive patients without clinical synovitis.* Ann Rheum Dis, 2016.
238. Gent, Y.Y.J., et al., *Macrophage positron emission tomography imaging as a biomarker for preclinical rheumatoid arthritis: Findings of a prospective pilot study.* Arthritis and Rheumatism, 2012. **64**(1): p. 62-66.
239. Kleyer, A., et al., *Bone loss before the clinical onset of rheumatoid arthritis in subjects with anticitrullinated protein antibodies.* Ann Rheum Dis, 2014. **73**(5): p. 854-60.
240. De Hair, M.J.H., et al., *Smoking and overweight determine the likelihood of developing rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2013. **72**(10): p. 1654-1658.
241. van de Stadt, L.A. and D. van Schaardenburg, *Alcohol consumption protects against arthritis development in seropositive arthralgia patients.* Ann Rheum Dis, 2012. **71**(8): p. 1431-2.
242. Jiang, X., et al., *Anti-CarP antibodies in two large cohorts of patients with rheumatoid arthritis and their relationship to genetic risk factors, cigarette smoking and other autoantibodies.* Ann Rheum Dis, 2014. **73**(10): p. 1761-8.
243. Shi, J., et al., *Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(42): p. 17372-7.
244. Hunt, L., et al., *AB0259 Anti-Carbamylated Protein (ANTI-CARP) Antibodies Are Present in the Sera of Individuals at Different Stages of the Inflammatory Arthritis Continuum: Table 1.* Annals of the Rheumatic Diseases, 2014. **73**(Suppl 2): p. 889.3-890.
245. Shi, J., et al., *Anti-carbamylated protein (anti-CarP) antibodies precede the onset of rheumatoid arthritis.* Ann Rheum Dis, 2014. **73**(4): p. 780-3.
246. Goeb, V., et al., *Potential role of arthroscopy in the management of inflammatory arthritis.* Clin Exp Rheumatol, 2012. **30**(3): p. 429-35.
247. Lawson, C.A., et al., *Early rheumatoid arthritis is associated with a deficit in the CD4+CD25high regulatory T cell population in peripheral blood.* Rheumatology, 2006. **45**(10): p. 1210-7.
248. Burgoyne, C.H., et al., *Abnormal T cell differentiation persists in patients with rheumatoid arthritis in clinical remission and predicts relapse.* Annals of the Rheumatic Diseases, 2008. **67**(6): p. 750-7.
249. Saleem, B., et al., *Patients with RA in remission on TNF blockers: when and in whom can TNF blocker therapy be stopped?* Annals of the Rheumatic Diseases, 2010. **69**(9): p. 1636-42.
250. Ponchel, F., et al., *An immunological biomarker to predict MTX response in early RA.* Ann Rheum Dis, 2014. **73**(11): p. 2047-53.
251. Filkova, M., et al., *Association of circulating miR-223 and miR-16 with disease activity in patients with early rheumatoid arthritis.* Ann Rheum Dis, 2013.

252. Murata, K., et al., *Comprehensive microRNA Analysis Identifies miR-24 and miR-125a-5p as Plasma Biomarkers for Rheumatoid Arthritis.* PLoS ONE, 2013. **8**(7).
253. Wu, X., et al., *MicroRNA expression signatures during malignant progression from Barrett's esophagus to esophageal adenocarcinoma.* Cancer Prev Res (Phila), 2013. **6**(3): p. 196-205.
254. Orr, C., et al., *Synovial tissue research: a state-of-the-art review.* Nat Rev Rheumatol, 2017. **13**(10): p. 630.
255. van de Sande, M.G., et al., *Local synovial engagement of angiogenic TIE-2 is associated with the development of persistent erosive rheumatoid arthritis in patients with early arthritis.* Arthritis Rheum, 2013. **65**(12): p. 3073-83.
256. Kraan, M.C., et al., *Immunohistological analysis of synovial tissue for differential diagnosis in early arthritis.* Rheumatology (Oxford), 1999. **38**(11): p. 1074-80.
257. Dennis, G., Jr., et al., *Synovial phenotypes in rheumatoid arthritis correlate with response to biologic therapeutics.* Arthritis Res Ther, 2014. **16**(2): p. R90.
258. van de Sande, M.G., et al., *Evaluating antirheumatic treatments using synovial biopsy: a recommendation for standardisation to be used in clinical trials.* Ann Rheum Dis, 2011. **70**(3): p. 423-7.
259. Lazarou, I., et al., *Ultrasound-guided synovial biopsy: a systematic review according to the OMERACT filter and recommendations for minimal reporting standards in clinical studies.* Rheumatology (Oxford), 2015. **54**(10): p. 1867-75.
260. Sanati, G., et al., *Innate lymphoid cells are pivotal actors in allergic, inflammatory and autoimmune diseases.* Expert Rev Clin Immunol, 2015. **11**(8): p. 885-95.
261. Spits, H., et al., *Innate lymphoid cells--a proposal for uniform nomenclature.* Nat Rev Immunol, 2013. **13**(2): p. 145-9.
262. Thabet, M.M., et al., *The prognostic value of baseline erosions in undifferentiated arthritis.* Arthritis Res Ther, 2009. **11**(5): p. R155.
263. Jansen, L.M., et al., *Predictors of radiographic joint damage in patients with early rheumatoid arthritis.* Ann Rheum Dis, 2001. **60**(10): p. 924-7.
264. Brown, A.K., *Using ultrasonography to facilitate best practice in diagnosis and management of RA.* Nat Rev Rheumatol, 2009. **5**(12): p. 698-706.
265. Backhaus, M., et al., *Guidelines for musculoskeletal ultrasound in rheumatology.* Ann Rheum Dis, 2001. **60**(7): p. 641-9.
266. Brown, A.K., et al., *An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis.* Arthritis Rheum, 2008. **58**(10): p. 2958-67.
267. Salaffi, F., et al., *Inter-observer agreement of standard joint counts in early rheumatoid arthritis: a comparison with grey scale ultrasonography--a preliminary study.* Rheumatology (Oxford), 2008. **47**(1): p. 54-8.
268. van de Stadt, L.A., et al., *The value of ultrasonography in predicting arthritis in auto-antibody positive arthralgia patients: a prospective cohort study.* Arthritis Res Ther, 2010. **12**(3): p. R98.
269. Colebatch, A.N., et al., *EULAR recommendations for the use of imaging of the joints in the clinical management of rheumatoid arthritis.* Ann Rheum Dis, 2013. **72**(6): p. 804-14.
270. Kraan, M.C., et al., *Asymptomatic synovitis precedes clinically manifest arthritis.* Arthritis Rheum, 1998. **41**(8): p. 1481-8.

271. Gent, Y.Y., et al., *Subclinical synovitis detected by macrophage PET, but not MRI, is related to short-term flare of clinical disease activity in early RA patients: an exploratory study.* Arthritis Res Ther, 2015. **17**: p. 266.
272. Gent, Y.Y., et al., *Macrophage positron emission tomography imaging as a biomarker for preclinical rheumatoid arthritis: findings of a prospective pilot study.* Arthritis Rheum, 2012. **64**(1): p. 62-6.
273. Mustila, A., et al., *Anti-citrullinated peptide antibodies and the progression of radiographic joint erosions in patients with early rheumatoid arthritis treated with FIN-RACo combination and single disease-modifying antirheumatic drug strategies.* Clin Exp Rheumatol, 2011. **29**(3): p. 500-5.
274. Syversen, S.W., et al., *Prediction of radiographic progression in rheumatoid arthritis and the role of antibodies against mutated citrullinated vimentin: results from a 10-year prospective study.* Ann Rheum Dis, 2010. **69**(2): p. 345-51.
275. van Steenbergen, H.W., et al., *The effects of rheumatoid factor and anticitrullinated peptide antibodies on bone erosions in rheumatoid arthritis.* Ann Rheum Dis, 2015. **74**(1): p. e3.
276. Hecht, C., et al., *Additive effect of anti-citrullinated protein antibodies and rheumatoid factor on bone erosions in patients with RA.* Ann Rheum Dis, 2015. **74**(12): p. 2151-6.
277. Lu, B., et al., *Being overweight or obese and risk of developing rheumatoid arthritis among women: a prospective cohort study.* Ann Rheum Dis, 2014. **73**(11): p. 1914-22.
278. de Smit, M., et al., *Antibodies against Porphyromonas gingivalis in seropositive arthralgia patients do not predict development of rheumatoid arthritis.* Ann Rheum Dis, 2014. **73**(6): p. 1277-9.
279. Johansson, L., et al., *Concentration of antibodies against Porphyromonas gingivalis is increased before the onset of symptoms of rheumatoid arthritis.* Arthritis Res Ther, 2016. **18**: p. 201.
280. Kharlamova, N., et al., *Antibodies to Porphyromonas gingivalis Indicate Interaction Between Oral Infection, Smoking, and Risk Genes in Rheumatoid Arthritis Etiology.* Arthritis Rheumatol, 2016. **68**(3): p. 604-13.
281. Brown, L.J. and H. Loe, *Prevalence, extent, severity and progression of periodontal disease.* Periodontol 2000, 1993. **2**: p. 57-71.
282. Quirke, A.M., et al., *Heightened immune response to autocitrullinated Porphyromonas gingivalis peptidylarginine deiminase: a potential mechanism for breaching immunologic tolerance in rheumatoid arthritis.* Ann Rheum Dis, 2014. **73**(1): p. 263-9.
283. Mankia, K., et al., *OP0246 Increased Prevalence of Periodontal Disease in Anti-CCP Positive Individuals at Risk of Progression To Inflammatory Arthritis: A Target for Prevention?* Annals of the Rheumatic Diseases, 2016. **75**(Suppl 2): p. 151.
284. Jin, Z., et al., *Alcohol consumption as a preventive factor for developing rheumatoid arthritis: a dose-response meta-analysis of prospective studies.* Ann Rheum Dis, 2014. **73**(11): p. 1962-7.
285. Isaacs, J.D. and G. Ferraccioli, *The need for personalised medicine for rheumatoid arthritis.* Ann Rheum Dis, 2011. **70**(1): p. 4-7.
286. Visser, H., et al., *How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis.* Arthritis & Rheumatism, 2002. **46**(2): p. 357-65.

287. Burgers, L.E., et al., *Clinical trials aiming to prevent Rheumatoid Arthritis cannot detect prevention without adequate risk stratification; the PROMPT-trial as example.* Arthritis Rheumatol, 2017.
288. van de Stadt, L.A., et al., *A prediction rule for the development of arthritis in seropositive arthralgia patients.* Ann Rheum Dis, 2013. **72**(12): p. 1920-6.
289. Karlson, E.W., et al., *Vitamin E in the primary prevention of rheumatoid arthritis: The women's health study.* Arthritis Care and Research, 2008. **59**(11): p. 1589-1595.
290. Walitt, B., et al., *Effects of postmenopausal hormone therapy on rheumatoid arthritis: the women's health initiative randomized controlled trials.* Arthritis and Rheumatism, 2008. **59**(3): p. 302-10.
291. Jick, S.S., et al., *Hyperlipidaemia, statin use and the risk of developing rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2009. **68**(4): p. 546-551.
292. Shadick, N.A., et al., *Low-dose aspirin in the primary prevention of rheumatoid arthritis: The women's health study.* Arthritis Care and Research, 2010. **62**(4): p. 545-550.
293. Bos, W.H., et al., *Effect of dexamethasone on autoantibody levels and arthritis development in patients with arthralgia: a randomised trial.* Ann Rheum Dis, 2010. **69**(3): p. 571-4.
294. Health, N.I.o. *Strategy to Prevent the Onset of Clinically-Apparent Rheumatoid Arthritis (StopRA) NCT02603146*. 2017 [cited 2017 26th March]; Available from: https://clinicaltrials.gov/ct2/show/NCT02603146.
295. Emery, P., et al., *Induction of Remission in Patients with up to 12 months of moderate-to-severe Rheumatoid Arthritis Symptoms treated with Etanercept Plus Methotrexate Over 52 Weeks. [abstract].* Arthritis & Rheumatism, 2012. **64 Suppl 10**: p. 2549.
296. Villeneuve, E., et al., *Preliminary Results of a Multicentre Randomised Control Trials of Etanercept and Methotrexate to Induce Remission in Patients with Newly Diagnosed Inflammatory Arthritis [Abstract].* Arthritis & Rheumatism, 2011. **63 Suppl 10**.
297. Tak, P.P., et al., *Sustained inhibition of progressive joint damage with rituximab plus methotrexate in early active rheumatoid arthritis: 2-year results from the randomised controlled trial IMAGE.* Ann Rheum Dis, 2012. **71**(3): p. 351-7.
298. Cohen, S.B., et al., *Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: Results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial evaluating primary efficacy and safety at twenty-four weeks.* Arthritis Rheum, 2006. **54**(9): p. 2793-806.
299. Westhovens, R., et al., *Clinical efficacy and safety of abatacept in methotrexate-naive patients with early rheumatoid arthritis and poor prognostic factors.* Ann Rheum Dis, 2009. **68**(12): p. 1870-7.
300. Genovese, M.C., et al., *Efficacy and safety of the selective co-stimulation modulator abatacept following 2 years of treatment in patients with rheumatoid arthritis and an inadequate response to anti-tumour necrosis factor therapy.* Ann Rheum Dis, 2008. **67**(4): p. 547-54.
301. Wakefield, R.J., et al., *Musculoskeletal ultrasound including definitions for ultrasonographic pathology.* J Rheumatol, 2005. **32**(12): p. 2485-7.
302. D'Agostino, M.A., R. Wakefield, and E. Filippucci, *Intra- and inter-observer reliability of ultrasonography for detecting and scoring synovitis*

in rheumatoid arthritis: a report of a EULAR ECSISIT task force. Ann Rheum Dis, 2005. **64(Suppl III)**: p. 62.
303. Ostergaard, M., et al., *An introduction to the EULAR-OMERACT rheumatoid arthritis MRI reference image atlas.* Ann Rheum Dis, 2005. **64 Suppl 1**: p. i3-7.
304. Glinatsi, D., et al., *Development and Validation of the OMERACT Rheumatoid Arthritis Magnetic Resonance Tenosynovitis Scoring System in a Multireader Exercise.* J Rheumatol, 2017.
305. Eshed, I., et al., *Tenosynovitis of the flexor tendons of the hand detected by MRI: an early indicator of rheumatoid arthritis.* Rheumatology (Oxford), 2009. **48**(8): p. 887-91.
306. Haavardsholm, E.A., et al., *Introduction of a novel magnetic resonance imaging tenosynovitis score for rheumatoid arthritis: reliability in a multireader longitudinal study.* Ann Rheum Dis, 2007. **66**(9): p. 1216-20.
307. McQueen, F.M., *The MRI view of synovitis and tenosynovitis in inflammatory arthritis: implications for diagnosis and management.* Ann N Y Acad Sci, 2009. **1154**: p. 21-34.
308. McGonagle, D., et al., *MRI for the assessment and monitoring of RA--what can it tell us?* Nat Rev Rheumatol, 2011. **7**(3): p. 185-9.
309. Hodgson, R., et al., *Dynamic contrast enhanced MRI of bone marrow oedema in rheumatoid arthritis.* Ann Rheum Dis, 2008. **67**(2): p. 270-2.
310. Nieuwenhuis, W.P., et al., *Evaluation of magnetic resonance imaging-detected tenosynovitis in the hand and wrist in early arthritis.* Arthritis Rheumatol, 2015. **67**(4): p. 869-76.
311. Brown, A.K., et al., *Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression.* Arthritis Rheum, 2006. **54**(12): p. 3761-73.
312. McQueen, F.M., et al., *Magnetic resonance imaging of the wrist in early rheumatoid arthritis reveals progression of erosions despite clinical improvement.* Ann Rheum Dis, 1999. **58**(3): p. 156-63.
313. Haavardsholm, E.A., et al., *Reliability and sensitivity to change of the OMERACT rheumatoid arthritis magnetic resonance imaging score in a multireader, longitudinal setting.* Arthritis Rheum, 2005. **52**(12): p. 3860-7.
314. McGonagle, D., et al., *The relationship between synovitis and bone changes in early untreated rheumatoid arthritis: a controlled magnetic resonance imaging study.* Arthritis Rheum, 1999. **42**(8): p. 1706-11.
315. Navalho, M., et al., *Bilateral MR imaging of the hand and wrist in early and very early inflammatory arthritis: tenosynovitis is associated with progression to rheumatoid arthritis.* Radiology, 2012. **264**(3): p. 823-33.
316. McQueen, F.M., et al., *Bone edema scored on magnetic resonance imaging scans of the dominant carpus at presentation predicts radiographic joint damage of the hands and feet six years later in patients with rheumatoid arthritis.* Arthritis Rheum, 2003. **48**(7): p. 1814-27.
317. Haavardsholm, E.A., et al., *Magnetic resonance imaging findings in 84 patients with early rheumatoid arthritis: bone marrow oedema predicts erosive progression.* Ann Rheum Dis, 2008. **67**(6): p. 794-800.
318. Krabben, A., et al., *MRI-detected subclinical joint inflammation is associated with radiographic progression.* Ann Rheum Dis, 2014. **73**(11): p. 2034-7.

319. Baker, J.F., et al., *Early MRI measures independently predict 1-year and 2-year radiographic progression in rheumatoid arthritis: secondary analysis from a large clinical trial.* Ann Rheum Dis, 2014. **73**(11): p. 1968-74.
320. Herz, B., et al., *Osteitis and synovitis, but not bone erosion, is associated with proteoglycan loss and microstructure damage in the cartilage of patients with rheumatoid arthritis.* Ann Rheum Dis, 2014. **73**(6): p. 1101-6.
321. Schleich, C., et al., *Intra-individual assessment of inflammatory severity and cartilage composition of finger joints in rheumatoid arthritis.* Skeletal Radiol, 2015. **44**(4): p. 513-8.
322. Boeters, D.M., et al., *MRI-detected osteitis is not associated with the presence or level of ACPA alone, but with the combined presence of ACPA and RF.* Arthritis Res Ther, 2016. **18**: p. 179.
323. Stomp, W., et al., *Are rheumatoid arthritis patients discernible from other early arthritis patients using 1.5T extremity magnetic resonance imaging? a large cross-sectional study.* J Rheumatol, 2014. **41**(8): p. 1630-7.
324. van Steenbergen, H.W., et al., *Characterising arthralgia in the preclinical phase of rheumatoid arthritis using MRI.* Ann Rheum Dis, 2014.
325. van Steenbergen, H.W., et al., *Clinical factors, anticitrullinated peptide antibodies and MRI-detected subclinical inflammation in relation to progression from clinically suspect arthralgia to arthritis.* Ann Rheum Dis, 2015.
326. Machado, P.M., et al., *The value of magnetic resonance imaging and ultrasound in undifferentiated arthritis: a systematic review.* J Rheumatol Suppl, 2011. **87**: p. 31-7.
327. Duer, A., et al., *Magnetic resonance imaging and bone scintigraphy in the differential diagnosis of unclassified arthritis.* Ann Rheum Dis, 2008. **67**(1): p. 48-51.
328. Duer-Jensen, A., et al., *Bone edema on magnetic resonance imaging is an independent predictor of rheumatoid arthritis development in patients with early undifferentiated arthritis.* Arthritis Rheum, 2011. **63**(8): p. 2192-202.
329. Tamai, M., et al., *A prediction rule for disease outcome in patients with undifferentiated arthritis using magnetic resonance imaging of the wrists and finger joints and serologic autoantibodies.* Arthritis Rheum, 2009. **61**(6): p. 772-8.
330. Ji, L., et al., *Early prediction of rheumatoid arthritis by magnetic resonance imaging in the absence of anti-cyclic citrullinated peptide antibodies and radiographic erosions in undifferentiated inflammatory arthritis patients: a prospective study.* Int J Rheum Dis, 2015. **18**(8): p. 859-65.
331. Nieuwenhuis, W.P., et al., *Older age is associated with more MRI-detected inflammation in hand and foot joints.* Rheumatology (Oxford), 2016. **55**(12): p. 2212-2219.
332. Mangnus, L., et al., *OP0218 Prevalence of MRI-Detected Inflammation in Symptom-Free Persons from the General Population and the Generation of Age-Dependent Ramris-Based Reference Values.* Annals of the Rheumatic Diseases, 2015. **74**(Suppl 2): p. 153.
333. Mangnus, L., et al., *MR-detected features of inflammation and erosions occur in symptom-free persons from the general population.* Arthritis Rheumatol, 2016.

334. Ostergaard, M., et al., *Magnetic resonance imaging-determined synovial membrane volume as a marker of disease activity and a predictor of progressive joint destruction in the wrists of patients with rheumatoid arthritis.* Arthritis Rheum, 1999. **42**(5): p. 918-29.
335. Ostergaard, M., et al., *New radiographic bone erosions in the wrists of patients with rheumatoid arthritis are detectable with magnetic resonance imaging a median of two years earlier.* Arthritis Rheum, 2003. **48**(8): p. 2128-31.
336. McQueen, F.M., *Magnetic resonance imaging in early inflammatory arthritis: what is its role?* Rheumatology (Oxford), 2000. **39**(7): p. 700-6.
337. Huang, J., et al., *A 1-year follow-up study of dynamic magnetic resonance imaging in early rheumatoid arthritis reveals synovitis to be increased in shared epitope-positive patients and predictive of erosions at 1 year.* Rheumatology (Oxford), 2000. **39**(4): p. 407-16.
338. Wakefield, R.J., et al., *Finger tendon disease in untreated early rheumatoid arthritis: a comparison of ultrasound and magnetic resonance imaging.* Arthritis Rheum, 2007. **57**(7): p. 1158-64.
339. Mangnus, L., et al., *Bone mineral density loss in clinically suspect arthralgia is associated with subclinical inflammation and progression to clinical arthritis.* Scand J Rheumatol, 2017: p. 1-5.
340. Buisman, L.R., et al., *A five-year model to assess the early cost-effectiveness of new diagnostic tests in the early diagnosis of rheumatoid arthritis.* Arthritis Res Ther, 2016. **18**(1): p. 135.
341. Romaine, S.P.R., et al., *MicroRNAs in cardiovascular disease: an introduction for clinicians.* Heart, 2015.
342. Kozomara, A. and S. Griffiths-Jones, *miRBase: annotating high confidence microRNAs using deep sequencing data.* Nucleic Acids Research, 2014. **42**(D1): p. D68-D73.
343. Yamada, H., et al., *Circulating microRNAs in autoimmune thyroid diseases.* Clin Endocrinol (Oxf), 2014. **81**(2): p. 276-81.
344. O'Connell, R.M., et al., *Physiological and pathological roles for microRNAs in the immune system.* Nat Rev Immunol, 2010. **10**(2): p. 111-122.
345. Carlsen, A.L., et al., *Circulating microRNA expression profiles associated with systemic lupus erythematosus.* Arthritis Rheum, 2013. **65**(5): p. 1324-34.
346. Baltimore, D., et al., *MicroRNAs: new regulators of immune cell development and function.* Nat Immunol, 2008. **9**(8): p. 839-845.
347. Kinoshita, T., et al., *MicroRNAs in extracellular vesicles: potential cancer biomarkers.* J Hum Genet, 2016.
348. Pichler, M. and G.A. Calin, *MicroRNAs in cancer: from developmental genes in worms to their clinical application in patients.* Br J Cancer, 2015. **113**(4): p. 569-573.
349. Spizzo, R., et al., *SnapShot: MicroRNAs in Cancer.* Cell, 2009. **137**(3): p. 586-586.e1.
350. Alevizos, I. and G.G. Illei, *MicroRNAs as biomarkers in rheumatic diseases.* Nature Reviews Rheumatology, 2010. **6**(7): p. 391-398.
351. Mitchell, P.S., et al., *Circulating microRNAs as stable blood-based markers for cancer detection.* Proc Natl Acad Sci U S A, 2008. **105**(30): p. 10513-8.

352. Chen, X., et al., *Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases.* Cell Res, 2008. **18**(10): p. 997-1006.
353. Doleshal, M., et al., *Evaluation and Validation of Total RNA Extraction Methods for MicroRNA Expression Analyses in Formalin-Fixed, Paraffin-Embedded Tissues.* The Journal of Molecular Diagnostics, 2008. **10**(3): p. 203-211.
354. Zhang, X., et al., *An Array-Based Analysis of MicroRNA Expression Comparing Matched Frozen and Formalin-Fixed Paraffin-Embedded Human Tissue Samples.* The Journal of Molecular Diagnostics, 2008. **10**(6): p. 513-519.
355. Wittmann, J. and H.-M. Jäck, *microRNAs in rheumatoid arthritis: midget RNAs with a giant impact.* Annals of the Rheumatic Diseases, 2011. **70**(Suppl 1): p. i92-i96.
356. Pauley, K.M., et al., *Upregulated miR-146a expression in peripheral blood mononuclear cells from rheumatoid arthritis patients.* Arthritis Res Ther, 2008. **10**(4): p. R101.
357. Murata, K., et al., *Plasma and synovial fluid microRNAs as potential biomarkers of rheumatoid arthritis and osteoarthritis.* Arthritis Res Ther, 2010. **12**(3): p. R86.
358. Trenkmann, M., et al., *Tumor necrosis factor alpha-induced microRNA-18a activates rheumatoid arthritis synovial fibroblasts through a feedback loop in NF-kappaB signaling.* Arthritis Rheum, 2013. **65**(4): p. 916-27.
359. Dong, L., et al., *Decreased expression of microRNA-21 correlates with the imbalance of Th17 and Treg cells in patients with rheumatoid arthritis.* J Cell Mol Med, 2014. **18**(11): p. 2213-24.
360. Lin, J., et al., *A novel p53/microRNA-22/Cyr61 axis in synovial cells regulates inflammation in rheumatoid arthritis.* Arthritis Rheumatol, 2014. **66**(1): p. 49-59.
361. Niederer, F., et al., *Down-regulation of microRNA-34a* in rheumatoid arthritis synovial fibroblasts promotes apoptosis resistance.* Arthritis Rheum, 2012. **64**(6): p. 1771-9.
362. Zhou, Q., et al., *Decreased expression of miR-146a and miR-155 contributes to an abnormal Treg phenotype in patients with rheumatoid arthritis.* Ann Rheum Dis, 2015. **74**(6): p. 1265-74.
363. Kurowska-Stolarska, M., et al., *MicroRNA-155 as a proinflammatory regulator in clinical and experimental arthritis.* Proc Natl Acad Sci U S A, 2011. **108**(27): p. 11193-8.
364. Fulci, V., et al., *miR-223 is overexpressed in T-lymphocytes of patients affected by rheumatoid arthritis.* Hum Immunol, 2010. **71**(2): p. 206-11.
365. Shibuya, H., et al., *Overexpression of microRNA-223 in rheumatoid arthritis synovium controls osteoclast differentiation.* Mod Rheumatol, 2013. **23**(4): p. 674-85.
366. Krintel, S., et al., *Prediction of treatment response to adalimumab: a double-blind placebo-controlled study of circulating microRNA in patients with early rheumatoid arthritis.* The pharmacogenomics journal, 2015.
367. Andersen, C.L., J.L. Jensen, and T.F. Orntoft, *Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets.* Cancer Res, 2004. **64**(15): p. 5245-50.

368. McCall, M.N., et al., *On non-detects in qPCR data.* Bioinformatics, 2014. **30**(16): p. 2310-6.
369. Julious, S.A., *Sample size of 12 per group rule of thumb for a pilot study.* Pharmaceutical Statistics, 2005. **4**(4): p. 287-291.
370. Tarn, J.C., S.; Gillespie, C.; Al-Ali, S.; James, K.; Locke, J.; Bowman, S.; Griffiths, B.; Young, D.; Ng, WF. *Whole Blood microRNA Signature for Primary Sjogren's Syndrome-Related Lymphoma (522)*. in *ACR/ARHP Annual Meeting*. 2014. Boston, MA, USA: Wiley-Blackwell Publishing.
371. Xiong, J., *Emerging roles of microRNA-22 in human disease and normal physiology.* Curr Mol Med, 2012. **12**(3): p. 247-58.
372. Xu, D., et al., *miR-22 represses cancer progression by inducing cellular senescence.* J Cell Biol, 2011. **193**(2): p. 409-24.
373. Zhang, J., et al., *microRNA-22, downregulated in hepatocellular carcinoma and correlated with prognosis, suppresses cell proliferation and tumourigenicity.* Br J Cancer, 2010. **103**(8): p. 1215-20.
374. Xia, S.S., et al., *MicroRNA-22 suppresses the growth, migration and invasion of colorectal cancer cells through a Sp1 negative feedback loop.* Oncotarget, 2017.
375. Jiang, R., et al., *miR-22 promotes HBV-related hepatocellular carcinoma development in males.* Clin Cancer Res, 2011. **17**(17): p. 5593-603.
376. Tsuchiya, N., et al., *Tumor suppressor miR-22 determines p53-dependent cellular fate through post-transcriptional regulation of p21.* Cancer Res, 2011. **71**(13): p. 4628-39.
377. Firestein, G.S., et al., *Apoptosis in rheumatoid arthritis: p53 overexpression in rheumatoid arthritis synovium.* The American journal of pathology, 1996. **149**(6): p. 2143.
378. Inazuka, M., et al., *Analysis of p53 tumour suppressor gene somatic mutations in rheumatoid arthritis synovium.* Rheumatology (Oxford), 2000. **39**(3): p. 262-6.
379. Lin, J., et al., *Cyr61 Induces IL-6 Production by Fibroblast-like Synoviocytes Promoting Th17 Differentiation in Rheumatoid Arthritis.* The Journal of Immunology, 2012. **188**(11): p. 5776-5784.
380. Zhang, Q., et al., *A critical role of Cyr61 in interleukin-17-dependent proliferation of fibroblast-like synoviocytes in rheumatoid arthritis.* Arthritis Rheum, 2009. **60**(12): p. 3602-12.
381. Lin, J., et al., *A Novel p53/microRNA-22/Cyr61 Axis in Synovial Cells Regulates Inflammation in Rheumatoid Arthritis.* Arthritis & Rheumatology, 2014. **66**(1): p. 49-59.
382. Hunt, L., et al., *T cell subsets: an immunological biomarker to predict progression to clinical arthritis in ACPA-positive individuals.* Annals of the Rheumatic Diseases, 2015: p. annrheumdis-2015-207991.
383. Lanchbury, J.S. and C. Pitzalis, *Cellular immune mechanisms in rheumatoid arthritis and other inflammatory arthritides.* Curr Opin Immunol, 1993. **5**(6): p. 918-24.
384. McInnes, I.B. and G. Schett, *Cytokines in the pathogenesis of rheumatoid arthritis.* Nat Rev Immunol, 2007. **7**(6): p. 429-42.
385. Fox, D.A., *The role of T cells in the immunopathogenesis of rheumatoid arthritis: new perspectives.* Arthritis Rheum, 1997. **40**(4): p. 598-609.
386. Janossy, G., et al., *Rheumatoid arthritis: a disease of T-lymphocyte/macrophage immunoregulation.* Lancet, 1981. **2**(8251): p. 839-42.

387. Panayi, G.S., J.S. Lanchbury, and G.H. Kingsley, *The importance of the T cell in initiating and maintaining the chronic synovitis of rheumatoid arthritis.* Arthritis Rheum, 1992. **35**(7): p. 729-35.
388. Buch, M.H., E.M. Vital, and P. Emery, *Abatacept in the treatment of rheumatoid arthritis.* Arthritis Res Ther, 2008. **10 Suppl 1**: p. S5.
389. Buch, M.H., et al., *Mode of action of abatacept in rheumatoid arthritis patients having failed tumour necrosis factor blockade: a histological, gene expression and dynamic magnetic resonance imaging pilot study.* Ann Rheum Dis, 2009. **68**(7): p. 1220-7.
390. Firestein, G.S. and N.J. Zvaifler, *How important are T cells in chronic rheumatoid synovitis?: II. T cell-independent mechanisms from beginning to end.* Arthritis Rheum, 2002. **46**(2): p. 298-308.
391. Ponchel, F., et al., *CD4+ T-cell subsets in rheumatoid arthritis.* International Journal of Clinical Rheumatology, 2012. **7**(1): p. 37-53.
392. Sakaguchi, S., et al., *Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases.* J Immunol, 1995. **155**(3): p. 1151-64.
393. Morgan, M.E., et al., *CD25+ cell depletion hastens the onset of severe disease in collagen-induced arthritis.* Arthritis Rheum, 2003. **48**(5): p. 1452-60.
394. Cao, D., et al., *Isolation and functional characterization of regulatory CD25brightCD4+ T cells from the target organ of patients with rheumatoid arthritis.* Eur J Immunol, 2003. **33**(1): p. 215-23.
395. Mottonen, M., et al., *CD4+ CD25+ T cells with the phenotypic and functional characteristics of regulatory T cells are enriched in the synovial fluid of patients with rheumatoid arthritis.* Clin Exp Immunol, 2005. **140**(2): p. 360-7.
396. van Amelsfort, J.M., et al., *CD4(+)CD25(+) regulatory T cells in rheumatoid arthritis: differences in the presence, phenotype, and function between peripheral blood and synovial fluid.* Arthritis Rheum, 2004. **50**(9): p. 2775-85.
397. Liu, W., et al., *CD127 expression inversely correlates with FoxP3 and suppressive function of human CD4+ T reg cells.* J Exp Med, 2006. **203**(7): p. 1701-11.
398. Baecher-Allan, C., et al., *CD4+CD25high regulatory cells in human peripheral blood.* J Immunol, 2001. **167**(3): p. 1245-53.
399. Ponchel, F., et al., *Dysregulated lymphocyte proliferation and differentiation in patients with rheumatoid arthritis.* Blood, 2002. **100**(13): p. 4550-6.
400. Ponchel, F., et al., *Interleukin-7 deficiency in rheumatoid arthritis: consequences for therapy-induced lymphopenia.* Arthritis Res Ther, 2005. **7**(1): p. R80-92.
401. Unutmaz, D., P. Pileri, and S. Abrignani, *Antigen-independent activation of naive and memory resting T cells by a cytokine combination.* J Exp Med, 1994. **180**(3): p. 1159-64.
402. Unutmaz, D., F. Baldoni, and S. Abrignani, *Human naive T cells activated by cytokines differentiate into a split phenotype with functional features intermediate between naive and memory T cells.* Int Immunol, 1995. **7**(9): p. 1417-24.
403. Brennan, F.M., et al., *Resting CD4+ effector memory T cells are precursors of bystander-activated effectors: a surrogate model of*

rheumatoid arthritis synovial T-cell function. Arthritis Res Ther, 2008. **10**(2): p. R36.
404. Olerup, O. and H. Zetterquist, *HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation.* Tissue Antigens, 1992. **39**(5): p. 225-35.
405. Kallberg, H., et al., *Smoking is a major preventable risk factor for rheumatoid arthritis: estimations of risks after various exposures to cigarette smoke.* Ann Rheum Dis, 2011. **70**(3): p. 508-11.
406. Virtanen, A., V. Kairisto, and E. Uusipaikka, *Regression-based reference limits: determination of sufficient sample size.* Clin Chem, 1998. **44**(11): p. 2353-8.
407. Peduzzi, P., et al., *Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates.* J Clin Epidemiol, 1995. **48**(12): p. 1503-10.
408. Roos-Engstrand, E., et al., *Expansion of CD4(+)CD25(+)helper T cells without regulatory function in smoking and COPD.* Respir Res, 2011. **12**(1): p. 74.
409. Barcelo, B., et al., *Phenotypic characterisation of T-lymphocytes in COPD: abnormal CD4+CD25+ regulatory T-lymphocyte response to tobacco smoking.* Eur Respir J, 2008. **31**(3): p. 555-62.
410. Vittinghoff, E. and C.E. McCulloch, *Relaxing the rule of ten events per variable in logistic and Cox regression.* Am J Epidemiol, 2007. **165**(6): p. 710-8.
411. Chalan, P., et al., *Circulating CD4+CD161+ T lymphocytes are increased in seropositive arthralgia patients but decreased in patients with newly diagnosed rheumatoid arthritis.* PLoS One, 2013. **8**(11): p. e79370.
412. Stack, R.J., et al., *Symptom complexes in patients with seropositive arthralgia and in patients newly diagnosed with rheumatoid arthritis: a qualitative exploration of symptom development.* Rheumatology (Oxford), 2014. **53**(9): p. 1646-53.
413. van Tuyl, L.H., et al., *Impact of Symptoms on Daily Life in People at Risk of Rheumatoid Arthritis.* Musculoskeletal Care, 2016. **14**(3): p. 169-73.
414. van Steenbergen, H.W., et al., *EULAR definition of arthralgia suspicious for progression to rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2017. **76**(3): p. 491-496.
415. Burgers, L.E., et al., *Validation of the EULAR definition of arthralgia suspicious for progression to rheumatoid arthritis.* Rheumatology (Oxford), 2017.
416. Falahee, M., et al., *Patients' Perceptions of Their Relatives' Risk of Developing Rheumatoid Arthritis and of the Potential for Risk Communication, Prediction, and Modulation.* Arthritis Care Res (Hoboken), 2017. **69**(10): p. 1558-1565.
417. Stack, R.J., et al., *Perceptions of risk and predictive testing held by the first-degree relatives of patients with rheumatoid arthritis in England, Austria and Germany: a qualitative study.* BMJ Open, 2016. **6**(6): p. e010555.
418. Novotny, F., et al., *Primary prevention of rheumatoid arthritis: a qualitative study in a high-risk population.* Joint Bone Spine, 2013. **80**(6): p. 673-4.

419. Finckh, A., et al., *Preventive Treatments for Rheumatoid Arthritis: Issues Regarding Patient Preferences.* Curr Rheumatol Rep, 2016. **18**(8): p. 51.
420. Gerlag, D., et al., *OP0182 Prevention of Rheumatoid Arthritis by B Cell Directed Therapy in The Earliest Phase of The Disease: The Prairi Study.* Annals of the Rheumatic Diseases, 2016. **75**(Suppl 2): p. 125-126.
421. *Arthritis prevention in the pre-clinical phase of rheumatoid arthritis with abatacept.* . [cited 2017 05 October 2017]; Available from: https://www.clinicaltrialsregister.eu/ctr-search/trial/2013-001479-18/GB.
422. Kaplan, R., et al., *Evaluating many treatments and biomarkers in oncology: a new design.* J Clin Oncol, 2013. **31**(36): p. 4562-8.
423. Mandrekar, S.J. and D.J. Sargent, *Drug designs fulfilling the requirements of clinical trials aiming at personalizing medicine.* Chin Clin Oncol, 2014. **3**(2): p. 14.
424. *PRe-clinical EValuation of Novel Targets in RA* [cited 2017 07 October 2017]; Available from: http://www.preventra.net/.
425. Sparks, J.A., et al., *Personalized Risk Estimator for Rheumatoid Arthritis (PRE-RA) Family Study: rationale and design for a randomized controlled trial evaluating rheumatoid arthritis risk education to first-degree relatives.* Contemp Clin Trials, 2014. **39**(1): p. 145-57.
426. Sparks, J.A., et al., *Disclosure of personalized rheumatoid arthritis risk using genetics, biomarkers, and lifestyle factors to motivate health behavior improvements:A randomized controlled trial.* Arthritis Care Res (Hoboken), 2017.
427. McQueen, F., et al., *OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Summary of OMERACT 6 MR Imaging Module.* J Rheumatol, 2003. **30**(6): p. 1387-92.
428. Kessler, L.G., et al., *The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions.* Stat Methods Med Res, 2015. **24**(1): p. 9-26.
429. Hodgson, R.J., P. O'Connor, and R. Moots, *MRI of rheumatoid arthritis image quantitation for the assessment of disease activity, progression and response to therapy.* Rheumatology (Oxford), 2008. **47**(1): p. 13-21.
430. Ostergaard, M., *Different approaches to synovial membrane volume determination by magnetic resonance imaging: manual versus automated segmentation.* Br J Rheumatol, 1997. **36**(11): p. 1166-77.
431. Klarlund, M., et al., *Dynamic magnetic resonance imaging of the metacarpophalangeal joints in rheumatoid arthritis, early unclassified polyarthritis, and healthy controls.* Scand J Rheumatol, 2000. **29**(2): p. 108-15.
432. Gaffney, K., et al., *Quantitative assessment of the rheumatoid synovial microvascular bed by gadolinium-DTPA enhanced magnetic resonance imaging.* Ann Rheum Dis, 1998. **57**(3): p. 152-7.
433. Gaffney, K., et al., *Quantification of rheumatoid synovitis by magnetic resonance imaging.* Arthritis Rheum, 1995. **38**(11): p. 1610-7.
434. Waterton, J.C., et al., *Repeatability and response to therapy of dynamic contrast-enhanced magnetic resonance imaging biomarkers in rheumatoid arthritis in a large multicentre trial setting.* Eur Radiol, 2017. **27**(9): p. 3662-3668.
435. Cimmino, M.A., et al., *Dynamic gadolinium-enhanced magnetic resonance imaging of the wrist in patients with rheumatoid arthritis can*

discriminate active from inactive disease. Arthritis Rheum, 2003. **48**(5): p. 1207-13.
436. Konig, H., J. Sieper, and K.J. Wolf, *Rheumatoid arthritis: evaluation of hypervascular and fibrous pannus with dynamic MR imaging enhanced with Gd-DTPA.* Radiology, 1990. **176**(2): p. 473-7.
437. Ostergaard, M., I. Lorenzen, and O. Henriksen, *Dynamic gadolinium-enhanced MR imaging in active and inactive immunoinflammatory gonarthritis.* Acta Radiol, 1994. **35**(3): p. 275-81.
438. Szkudlarek, M., et al., *Power Doppler ultrasonography for assessment of synovitis in the metacarpophalangeal joints of patients with rheumatoid arthritis: a comparison with dynamic magnetic resonance imaging.* Arthritis Rheum, 2001. **44**(9): p. 2018-23.
439. Szkudlarek, M., et al., *Contrast-enhanced power Doppler ultrasonography of the metacarpophalangeal joints in rheumatoid arthritis.* Eur Radiol, 2003. **13**(1): p. 163-8.
440. Shen, H., J.C. Goodall, and J.S. Hill Gaston, *Frequency and phenotype of peripheral blood Th17 cells in ankylosing spondylitis and rheumatoid arthritis.* Arthritis Rheum, 2009. **60**(6): p. 1647-56.
441. Harrington, L.E., et al., *Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages.* Nature immunology, 2005. **6**(11): p. 1123-32.
442. Park, H., et al., *A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17.* Nature immunology, 2005. **6**(11): p. 1133-41.
443. Leipe, J., et al., *Role of Th17 cells in human autoimmune arthritis.* Arthritis and rheumatism, 2010. **62**(10): p. 2876-85.
444. van Hamburg, J.P., et al., *Th17 cells, but not Th1 cells, from patients with early rheumatoid arthritis are potent inducers of matrix metalloproteinases and proinflammatory cytokines upon synovial fibroblast interaction, including autocrine interleukin-17A production.* Arthritis and rheumatism, 2011. **63**(1): p. 73-83.
445. Jovanovic, D.V., et al., *IL-17 stimulates the production and expression of proinflammatory cytokines, IL-beta and TNF-alpha, by human macrophages.* J Immunol, 1998. **160**(7): p. 3513-21.
446. Kotake, S., et al., *IL-17 in synovial fluids from patients with rheumatoid arthritis is a potent stimulator of osteoclastogenesis.* J Clin Invest, 1999. **103**(9): p. 1345-52.