Validation of the 2010 American College of Rheumatology/ European League Against Rheumatism classification criteria for rheumatoid arthritis with special emphasis on the role of autoantibodies

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Medical and Human Sciences

2015

Jennifer Helen Humphreys

School of Medicine

Table of Contents

List of tables

List of figures

Abstract

The University of Manchester

Jennifer Helen Humphreys

Doctor of Philosophy

Validation of the 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis with a special emphasis on the role of autoantibodies

April 2015

Aim: The aim of this thesis was to validate the 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis (RA), in particular with respect to its construct validity and the role of autoantibodies within the criteria.

Methods: This thesis used data from the Norfolk Arthritis Register, a longitudinal inception cohort of adults (≥16 years) with inflammatory polyarthritis (IP), defined as ≥ 2 swollen joint for ≥ 4 weeks. The 2010 criteria were used to define RA, firstly in a re-estimation of the incidence rates (IR) with comparisons made to incidence defined by the previous criteria set; and secondly in a study comparing mortality rates in patients with RA to those of the general population, and how these rates changed over time. Analyses were performed testing the ability of the 2010 criteria to identify those patients with IP at increased risk of mortality, disability, disease severity and radiographic damage. The levels and number of autoantibodies present were investigated as predictors of mortality in patients with IP. The association between anti-carbamylated protein (anti-CarP) antibodies and long term disease outcomes were investigated.

Results: The incidence of RA was 40 per 100 000 population; baseline IRs were similar to the cumulative IRs using the previous criteria set over 5 years. Patients who were seronegative were less likely to be classified as RA by the 2010 criteria. Mortality rates in patients with RA were higher compared to the general population (standardised mortality ratio 1.16, 95% confidence interval (CI) 1.04-1.29) and declined over the study period at the same rate as the general population. Patients with IP who fulfilled the 2010 criteria had increased risk of early death (hazard ratio (HR) 1.35, 95% CI 1.13-1.64), as well as increased levels of disability (β 0.38, 95% CI 0.33-0.43), disease severity (β 1.63, 95% CI 1.54-1.73) and radiographic damage (β 0.33, 95% CI 0.20-0.47) throughout follow up. Patients with two autoantibodies had an increased risk of early death (HR 1.35, 95% CI 1.09-1.68), but there was no association with early death and the levels of these antibodies. Anti-CarP antibody positivity was independently associated with worse disability (β 0.12, 95% CI 0.02-0.21) and disease severity (β 0.23, 95% CI 0.07-0.39) throughout follow up.

Conclusions: The 2010 ACR/EULAR classification criteria for RA identify patients with IP early in their disease course and recognise those at increased risk of mortality and poor outcomes. The 2010 criteria may miss a subgroup of seronegative patients who nevertheless have a poor prognosis. Novel autoantibodies may be useful to identify this subgroup.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

Acknowledgements

Firstly I would like to thank my supervisors, Professor Deborah Symmons, Dr Suzanne Verstappen and Dr Kimme Hyrich for their guidance, encouragement and wisdom throughout my PhD. Thank you to Anne Barton for her counsel as advisor. Thank you to all the members of the NOAR team in Manchester and Norfolk, in particular Jackie Chipping. I am grateful to the statisticians in the Arthritis Research UK Centre for Epidemiology, who have all, at various times throughout my PhD, provided essential advice and guidance. Thank you to all the NOAR nurses and patients who participated in the study, and in particular to Arthritis Research UK for funding this PhD. Special thanks to all the Clinical Research Fellows in the Centre for Musculoskeletal Research over the past 4 years, in particular John Reynolds, Sarah Skeoch and Meghna Jani for their friendship and objective honesty throughout, as well as providing as many cups of tea and coffee as required.

I would like to thank my family. Thank you to my mum and dad for a lifetime of support and encouragement. Finally, thank you to Gavin and Will. I would not have been able to complete this work without your love, patience and ability to make me laugh no matter how challenging things seemed.

Preface

I graduated from the University of Liverpool in 2006 with MBChB and an intercalated MSc Medical Microbiology (distinction). During my foundation years at Aintree University Hospital I developed an interest in Rheumatology and was encouraged to apply for one of the first Academic Foundation Posts, researching obesity in early arthritis. This confirmed my interest in pursuing a clinical academic career in Rheumatology and I then successfully applied for an NIHR Academic Clinical Fellowship at Keele University. Alongside this I progressed through my clinical training and during my first year as a Rheumatology Specialty Registrar I was awarded an Arthritis Research UK Clinical Research Fellowship at the University of Manchester. The work in this thesis is the product of that fellowship.

Role of the candidate in this PhD

My role encompassed all key aspects of the work presented in this thesis, including:

- Design and development of the research questions
- Application for funding
- Establishing collaborative projects
- Planning the analyses
- Coordinating extraction and transfer of stored serum samples for anti-carbamylated protein antibody testing
- Data entry for the anti-carbamylated protein antibody analysis
- Sourcing and coordinating access to and data entry of blood test results to supplement missing data
- Data cleaning
- Data preparation
- Statistical analysis
- Interpretation of results
- Presentation of findings at national and international conferences
- Writing of manuscripts submitted for publication
- Writing of this thesis

Rationale for alternative format

The alternative format was selected for this thesis because the nature of the PhD project fits better with the alternative format style. Instead of a single hypothesis with investigations along the way each building from the last to test this hypothesis, the project consisted of a series of discrete questions, each testing the validity of the 2010 American College of Rheumatology/European League against Rheumatism classification criteria for rheumatoid arthritis in different ways. In addition I have been successful in having a number of manuscripts based on my analyses accepted for publication. Along with my supervisory team it was therefore agreed that overall this project was more in keeping with this thesis format, and approval granted prospectively by the Chair of the Research Degrees Committee in the Faculty of Medical and Human Sciences at the University of Manchester.

The thesis has therefore been constructed as follows. The introduction and discussion are written in the same way as would be expected for a traditional format thesis. The methods chapter provides details on the methods of the Norfolk Arthritis Register (NOAR), which provided the data for this thesis, as well as the statistical methodologies employed in the analyses and the reasons for their selection. The first results chapter (chapter 3) outlines the baseline characteristics of all patients within NOAR, as well as retention in the register over time. The remaining results chapters comprise of 5 published manuscripts, grouped by topic, as well as one manuscript currently in submission. The analysis in section 6.1 addresses a key objective of this thesis, but is not currently planned for publication. It has been prepared in the paper style for consistency.

Publications

Humphreys JH, van Nies J, Chipping J, Marshall T, van der Helm-van Mil A, Symmons DP, Verstappen SM. Rheumatoid factor and anti-citrullinated protein antibody positivity, but not their level, are associated with increased mortality in patients with rheumatoid arthritis: results from two large independent cohorts. *Arthritis Research & Therapy* 2014; **16**:483

Humphreys JH, Warner A, Chipping J, Marshall T, Lunt M, Symmons DP, Verstappen SM. Mortality trends in patients with early rheumatoid arthritis over 20 years: Results from the Norfolk Arthritis Register. *Arthritis* Care & Research 2014; 66(9):1296-301

Humphreys JH, Verstappen SMM, Scire CA, Uhlig T, Fautrel B, Sokka T, Symmons DPM. How do we classify rheumatoid arthritis in established disease – Can we apply the 2010 ACR/EULAR classification criteria? (Viewpoint) *Journal of Rheumatology* 2014; 41(12):2347-2351.

Humphreys JH, Verstappen SMM. Etude et Suivie: Rheumatoid Arthritis in the 21st Century. (Editorial) *Journal of Rheumatology* 2013; 40(10):1637-9

Humphreys JH, Symmons DPM. Post publication validation of the 2010 ACR/EULAR classification criteria for rheumatoid Arthritis: where do we stand? (Review) *Current Opinion in Rheumatology* 2013; 25(2):157-63

Humphreys JH, Verstappen SM, Hyrich KL, Chipping JR, Symmons DP 2010 ACR/EULAR classification criteria for rheumatoid arthritis predict increased mortality in patients with early arthritis: results from the Norfolk Arthritis Register (NOAR). *Rheumatology* 2013; 52(6):1141-2

Humphreys JH, Verstappen SM, Hyrich KL, Chipping JR, Marshall T, Symmons DP. The incidence of rheumatoid arthritis in the UK: comparisons using the 2010 ACR/EULAR classification criteria and the 1987 ACR classification criteria. Results from the Norfolk Arthritis Register. *Annals of the Rheumatic Diseases* 2013; 72(8):1315-20

Humphreys J, Verstappen S, Mirjafari H, Bunn D, Lunt M, Bruce I, Symmons D. Morbid obesity is associated with disability in early inflammatory polyarthritis: Results from the Norfolk Arthritis Register (NOAR*). Arthritis Care & Research* 2013; 65(1):122-6

Abbreviations

ACPA	Anti-citrullinated protein antibody
ACR	American College of Rheumatology
Anti-CarP	Anti-carbamylated protein
Anti-CCP	Anti-cyclic citrullinated peptide
Anti-TNF	Anti-tumour necrosis factor alpha
APF	Anti-perinuclear factor
ARA	American Rheumatism Association
AR UK	Arthritis Research UK
AUC	Area under the (receiver operating characteristic) curve
CRP	C-reactive protein
CVD	Cardiovascular disease
DAG	Direct acyclic graph
DAS	Disease activity score
DAS28	Disease activity score for 28 joints
DMARD	Disease modifying anti-rheumatic drug
EAC	Leiden Early Arthritis Clinic
EDTA	Ethylenediaminetetraacetic acid
EIA	Early inflammatory arthritis
ESPOIR	Études et Suivi des Polyarthrites Indifférenciées Récentes
ESR	Erythrocyte sedimentation rate
ERAN	Early Rheumatoid Arthritis Network
ERAS	Early Rheumatoid Arthritis Study
EULAR	European League Against Rheumatism
GP	General practitioner
HAQ	Health assessment questionnaire
HLA-DR4	Human leukocyte antigen DR4
HR	Hazard ratio
HSCIC	Health and Social Car Information Centre
ICD	International Classification of Diseases
IQR	Interquartile range
IP	Inflammatory polyarthritis
IU	International units
LR+	Positive likelihood ratio
LR-	Negative likelihood ratio
MHC	Major histocompatibility complex

MRI	Magnetic resonance imaging
MRR	Mortality rate ratio
MTX	Methotrexate
NHS-IC	National Health Service Information Centre
NHSI	Nurses' Health Study I
NHSII	Nurses' Health Study II
NOAR	Norfolk Arthritis Register
NPV	Negative predictive value
NSAID	Non-steroidal anti-inflammatory drug
ONS	Office for National Statistics
PPV	Positive predictive value
PsA	Psoriatic arthritis
RA	Rheumatoid arthritis
RCT	Randomised controlled trial
RF	Rheumatoid factor
ROC	Receiver operating characteristic
SE	Shared epitope
SLE	Systemic lupus erythematosus
SMR	Standardised mortality ratios
UA	Undifferentiated arthritis
ULN	Upper limit of normal
US	Ultrasound

Introduction

This chapter gives an overview of the history of classification criteria in rheumatoid arthritis (RA) and discusses selected autoantibodies in RA. A systematic review appraises the literature on validation studies of the 2010 American College of Rheumatology/ European League Against Rheumatism classification criteria for RA. Finally the aims and objectives of this thesis are outlined.

1 Introduction

1.1 Rheumatoid Arthritis

Rheumatoid arthritis (RA) is a chronic inflammatory disease affecting approximately 1% of the adult population (1). It is associated with significant disability, morbidity and increased mortality (2-5). It is characterised by inflammation of the synovial joints, typically the small joints of the hands, and is associated with the presence of the autoantibodies rheumatoid factor (RF) and anti-citrullinated protein antibodies (ACPA) in the sera. Typical bony erosions and progressive destruction of the joints occurs if patients are untreated, which are both painful and disabling. The average age of onset is the 6th decade and there is a female preponderance with a ratio of 3 women affected for every man (1;6). RA is a multisystem disease not only confined to the joints; there are recognised associations with increased risk of cardiovascular disease (7) and cancer (8). In addition, RA patients may develop respiratory manifestations in the form of pulmonary nodules or fibrosis, and other rarer manifestations such as rheumatoid vasculitis, Felty's syndrome and amyloidosis (9). Many of these extra-articular manifestations are considered to be the consequence of a chronic inflammatory burden. Although recognised as an autoimmune disease, the specific aetiology of RA is unknown. There is a strong genetic component, with the heritability estimated to be up to 68% (10), however no single genetic trait has been recognised in all patients with RA. It therefore falls into the group of complex genetic disorders, alongside type 1 diabetes mellitus and asthma. In RA, as with other complex disorders, pathogenesis is thought to occur due to the interplay between genetic and environmental factors that trigger the onset of joint inflammation (11). Smoking is the most frequently recognised environmental factor associated with the disease, and has been shown to have a dose dependent relationship with risk of developing RA (12). Other environmental risk factors include obesity (13) and low socioeconomic status (12). Alcohol intake and breastfeeding may have a protective effect on the risk of developing RA (12;14;15). The relationship between these environmental risk factors and the various underlying genetic risk traits is complex and as yet not been fully described. In fact, it has been postulated that the clinical phenotype of RA is not one single disease but is the end product of many different pathogenetic pathways (16).

1.2 Clinical presentation and diagnosis

The classic description of RA is one of a symmetrical, deforming polyarthritis, associated with raised inflammatory markers, radiographic erosions and the presence in the sera of RF or ACPA. However in clinical practice, RA is a highly heterogeneous disease (2;17;18), and in its early stages it may be difficult of differentiate from a wide range of other inflammatory arthritides such as psoriatic arthritis (PsA), systemic lupus erythematosus (SLE) and reactive arthritis (13;19). Diagnosis is complicated by the lack of a single pathognomonic clinical, laboratory or radiological marker for RA. Further, even in patients with many typical features of RA, a proportion will go into spontaneous remission (20;21). In contrast, not all patients with RA will have the characteristic serological markers or a typical pattern of joint involvement, and those who present atypically may still evolve into severe disabling RA (18). Atypical or undifferentiated arthritis is common; many patients who subsequently satisfy classification criteria do not do so early in the disease course (17;22). Nevertheless, because of the difficulty in making a clinical diagnosis, classification criteria often guide diagnosis in clinical practice. Autoantibodies play a key role in classification criteria, as a diagnostic aide and a prognostic marker.

1.3 Autoantibodies in RA

There are two established autoantibodies associated with RA and with disease severity in the disease: RF and ACPA. These antibodies have been shown to be present before the onset of symptoms in cohorts of high risk asymptomatic individuals who later develop RA (23;24), and to predict development of synovitis in patients presenting with arthralgia but no active joint inflammation (25). Brink et al (26), in a study using biobank samples collected years prior to RA diagnosis, demonstrated an initial limited autoantibody response that widened markedly as time to diagnosis reduced. However, a number of patients lack either of these antibodies. It has yet to be established whether other, currently unidentified, antibodies are present in all of these patients or whether some are truly seronegative.

1.3.1 Rheumatoid factor

Rheumatoid factor is an autoantibody which develops against the Fc portion of immunoglobulin G (IgG) antibodies, in the presence of high inflammatory load and poor clearance of immune complexes (27). It was first described in autoimmune disease by Waaler in 1939 (28), in a single patient with rheumatoid arthritis. Subsequently, it was found to occur in the majority of patients with RA and therefore was adopted as an additional tool to aide diagnosis, and included in diagnostic and classification criteria (29-31). Over time, it was also shown to have prognostic properties. In particular it is associated with erosive disease (32;33), extra-articular manifestations (34;35), decreased chance of remission (36) and decreased survival (37;38).

Nevertheless, RF is an imperfect test to confirm RA. In particular, it is not specific to RA, having been demonstrated in the sera of patients with a range of autoimmune diseases, including Sjogren's syndrome, sarcoidosis and SLE (39). It has also been reported in chronic infections such as tuberculosis, infective endocarditis and leprosy (27). Most notably, perhaps, it can be seen in healthy individuals, with reported prevalence of RF in overall populations of 5-25%(40). The prevalence appears to have ethnic variability and is increased in older populations and smokers (27;40;41).

RF also has limitations as a marker of poor prognosis in patients with RA. The association between RF and levels of disease activity has not been consistently demonstrated (36;42), neither has an association with disability (36;43-45). As a result there remains a need to identify additional biomarkers both to aid diagnosis and provide clearer prognostic information.

1.3.2 Anti-citrullinated protein antibodies

ACPA are a group of autoantibodies directed against proteins that have undergone the post-translational modification called citrullination. This is an enzyme-driven conversion of arginine to citrulline that happens in the presence of peptidyl arginase (PAD4) (46). A proposed link between smoking exposure and ACPA positive disease is supported by evidence that smoking is a trigger of citrullination in the lungs (47).

The identification of ACPA in patients with RA was first shown in a study by Nienhuis who developed the anti-perinuclear factor (APF) test (48). However this was a complicated and cumbersome test and was never widely adopted. In 1995 it was found that the APF test was in fact detecting the presence of anti-filaggrin antibodies and more straightforward techniques for testing became available (49). ACPA are now most commonly tested for using a variety of commercially available anti-cyclic citrullinated peptide (anti-CCP) antibody tests, which are a panel of synthetically generated peptides recognised by ACPA (46). These are not the naturally occurring antigens and were developed because filaggrin is not present in the inflamed joint, so it was hypothesized that better antigenic peptides could be found. These anti-CCP antibodies were found to have similar sensitivity to RF, but also to be highly specific for RA; they are rarely identified in non-RA patients and specificity of up to 98% has been reported (50). This increased specificity has seen them rapidly adopted into clinical practice.

There is evidence from gene-environment studies to suggest that ACPA themselves have a pathogenic role. A combination of the presence of shared epitope (SE) genes, a region within the major histocompatibility complex (MHC) strongly associated with RA, and the environmental trigger of smoking may induce citrullination (47). They have identified a 21 fold increase in the risk of developing ACPA positive RA in smokers carrying two copies of the shared epitope. This is an interesting finding; however it should be remembered that even if such mechanisms can be confirmed, they can only inform causality in patients with ACPA positive disease and certain genotypes. Nevertheless, the presence of ACPA many years prior to the onset of symptoms also supports a potential pathogenic role for these antibodies. Further, an association has been reported between the presence of ACPA and serum markers of osteoclast-mediated bone resorption (51), potentially providing the mechanism for erosive disease in ACPA positive patients.

As a predictor of disease outcomes, ACPA have been shown to have associations with disease activity (52), increased mortality (53) and a particularly strong association with erosive damage to joints (54-56). They have also been shown to affect response to treatment, with ACPA positive patients responding less well to treatment than ACPA negative patients (57).

Despite the markedly improved specificity of ACPA over RF, in clinical practice we recognise there remains a subset of patients with inflammatory arthritis who lack either of these antibodies, but nevertheless have a poor prognosis. Differentiating

these patients from those whose disease course will be mild, and identifying them sufficiently early in their disease course when outcomes still have the potential to be modified by treatment, is a major challenge for clinicians and researchers.

1.3.3 Anti-carbamylated protein antibodies

Recently a novel family of autoantibodies has been identified which recognise proteins that have undergone a different port-translational modification, known as carbamylation (58). In carbamylation, lysine is converted to homocitrulline in the presence of citrate and urea. Like citrullination this can be a physiological process, but also like citrullination, because it produces proteins that are 'altered self', the development of autoantibodies may occur. As with ACPA and RF, these anti-carbamylated protein (anti-CarP) antibodies have been identified in the sera of asymptomatic blood donors who have later gone to develop RA (59), and predicted development of RA in a cohort of patients with arthralgia but no evidence of inflammatory arthritis (58). They are present in patients who are ACPA positive and negative, and in inhibition studies, have shown surprisingly little cross-reactivity with ACPA, despite the similarities of citrulline and homocitrulline (60).

Anti-CarP antibody status also shows potential as a prognostic biomarker in patients with inflammatory arthritis. In a study of patients with early inflammatory arthritis (EIA) from the Leiden Early Arthritis Clinic (EAC) anti-CarP antibodies were shown to be associated with erosive disease, and that association persisted when patients were subdivided by ACPA status (61). They are therefore a potential candidate for identifying patients we would currently consider to be seronegative who have poor prognosis. However there are currently no other data on associations between anti-CarP antibodies and long term outcomes such as disability or mortality. In the light of the 2010 criteria and their strong weighting towards ACPA/RF positivity, they may be of key importance.

1.4 Evolution of RA management over time

Historically, patients presenting with inflammatory arthritis were initially treated symptomatically with non-steroidal anti-inflammatory drugs (NSAIDs) and observed to allow natural differentiation into typical RA or resolution of symptoms. Only if there was evidence of persistent joint inflammation and progression of radiographic

changes were disease modifying anti-rheumatic drugs (DMARDs) introduced. DMARDs were preferably given as monotherapy and changed to alternatives if efficacy was not achieved or lost, rather than in given in combination. This strategy was known as the 'pyramid' approach to treatment, and aimed to minimise patients' exposure to potentially toxic and immunosuppressive therapies (62). However, as a result of this regime patients often had significant joint deformity and disability before disease modifying therapy was initiated (63-65).

A sequence of sometimes overlapping advances in rheumatology research and understanding led to a change in approach. In the late 1980s methotrexate was recognised as an effective treatment for RA and began to be used more widely (66). In the 1990s and early 2000s, a number of key studies identified firstly that better clinical outcomes are achieved the earlier treatment is started (67;68), and secondly that if treatment is started sufficiently early, it may be possible to switch off disease activity altogether (69;70). Alongside this, it became clear that toxicity from treatment with DMARDs, in particular methotrexate, was not inevitable and harm could be minimised with careful monitoring (71;72). Subsequently a number of studies also demonstrated marked benefits of treating patients with combinations of DMARDs, rather than sequential monotherapy (73;74). Finally, at the start of this century treatment advanced dramatically with the advent of the first group of biologic drugs to become widely used in RA, anti-tumour necrosis factor alpha (anti-TNF) therapies. This, along with the availability of multiple subsequent biologics, has provided effective treatments to patients with previously entirely refractory disease (75).

The inevitable consequence of this better understanding of the importance of early treatment and availability of a greater number of potential therapies has been a paradigm shift away from the 'pyramid approach' to a 'treat-to-target' approach. The aim of modern therapeutic strategy is to quickly achieve and maintain low disease activity or remission wherever possible. As these therapies and treatment strategies have evolved, it became clear that the available classification criteria for RA at the time were no longer fit for purpose, as they lacked sensitivity in early RA (76). Thus, classification criteria were required to evolve along with the management of the disease.

1.5 Classification Criteria in RA

RA is a purely clinical diagnosis, based on the judgment of a physician. In clinical practice this approach is reasonable, however in academic and research fields, it is essential to have a clear definition of the disease to allow for standardisation and comparisons between different studies; in order to do this, classification criteria have been developed. It is important to note that classification criteria and diagnostic criteria are not interchangeable terms, and that classification criteria should not be used in isolation to make a diagnosis; their purpose is to provide a case definition for research. Nevertheless, it is generally acknowledged that such criteria are often used to inform diagnostic decisions.

1.5.1 1958 revision of diagnostic criteria for RA

To date, there have been three sets of widely used classification criteria for RA. The first set of classification criteria for RA were proposed by the American Rheumatism Association (ARA) in 1956 (77) with minor revisions in 1958 (29) and were proposed as diagnostic criteria. Patients were divided into classical, definite, probable and possible RA, based on the presence of specific joint signs and symptoms (see figure 1.5.1). These first criteria provided the case definition of RA for nearly 30 years and were developed by a committee of 5 members of the ARA. The criteria were derived from previously published research (e.g. the Pittsburgh Arthritis Study (78)), and through the involvement of 'a number of physicians particularly interested in rheumatic diseases across the United States and Canada' who supplied case reports. As well as fulfilling the criteria presented in figure 1.5.1, there was also a large list of other diagnoses which had to first be excluded, such as SLE. However, over time it became clear that these criteria classified large numbers of patients in the community as having probable or definite RA whose disease was not progressive (79), which led to a misrepresentation within the medical community of RA as a benign and generally self-limiting disease. In addition, the criteria were thought to be too complex, the list of diagnostic exclusions too long for practical use; further, 3 of the criteria required invasive procedures (synovial biopsy, nodule biopsy and mucin clot) which were infrequently performed and often not feasible as part of routine clinical practice.

1958 revision of diagnostic criteria for RA

≥7 of the following criteria for ≥6 weeks = classical RA; ≥5 = definite RA; ≥3 = probable diagnosis of RA:

- Morning stiffness
- Pain on motion /tenderness in at least one joint (observed by physician)
- Swelling (not bony overgrowth alone) in at least one joint (observed by a physician)
- Swelling (observed by a physician) of at least one other joint (any interval free of joint symptoms between the 2 joint involvement may not be more than 3 months).
- Symmetrical joint swelling (observed by a physician) with simultaneous involvement of the same joint on both sides of the body (bilateral involvement of midphalangeal, metocarpophalangeal or metatarsophalangeal joints is acceptable without exact symmetry). Terminal phalangeal joint involvement will not satisfy this criterion.
- Subcutaneous nodules (observed by a physician)
- Radiographic changes typical of RA (which must include at least bony decalcification localized to or greatest around the involved joints, not just degenerative changes). Degenerative changes do not exclude patients from any group classified as RA.
- RF positive
- Poor mucin precipitate from synovial fluid (with shreds and cloudy solution).
- Characteristic histological change s in synovial membrane (≥3 of: marked villous hypertrophy; proliferation of superficial synovial cells often with palisading; marked infiltration of chronic inflammatory cells (lymphocytes or plasma cells predominating) with tendency to form 'lymphoid nodules'; deposition of compact fibrin, either on surface or interstitially; foci of cell necrosis)
- Characteristic histological changes in nodules (granulomatous foci with central zones of cell necrosis, surrounded by proliferated fixed cells, and peripheral fibrosis and chronic inflammatory cell infiltration, predominantly perivascular)

≥2 of the following criteria for ≥3 weeks = possible RA:

- Morning stiffness
- Tenderness or pain on motion (observed by a physician) with history of recurrence or persistence for 3 weeks
- History or observation of joint swelling.
- Subcutaneous nodules (observed by a physician).
- Elevated erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP)
- Iritis

Figure 1.5.1 The 1958 ARA diagnostic criteria for RA, adapted from (29)

1.5.2 ARA 1987 revised criteria for the classification of RA

In 1987 an attempt was made by the ARA to provide a more specific and less heterogeneous definition of RA (30), in order to allow clinicians and researchers to clearly differentiate between patients with RA and those with other conditions in which there may also be an inflammatory polyarthritis. The specific aim was to develop a classification criteria set, rather than diagnostic criteria. The development approach was systematic: 262 RA patients and 262 control subjects, who had all been diagnosed with a non-RA rheumatic disease such as osteoarthritis or SLE, were enrolled consecutively from the outpatient clinics of 41 rheumatologists. A list of potential disease discriminating features was suggested by the development committee using the Delphi process, and included the 1958 criteria set; data on these features were collected for all patients and controls. In addition, the diagnostic certainty of RA was assessed on a 10 cm visual analogue scale by the investigator. Univariate statistical analysis of each potential criterion was performed using chi-square tests, comparing cases and controls. Two methods were then used to develop the classification criteria. The combination of features found to be most sensitive and specific for RA was compiled into a traditional list format of 7 criteria (see figure 1.5.2). The second method created a classification tree, using the most discriminative variable to divide the participants into RA and non-RA, then repeating the procedure with the second most discriminative variable in the resulting subgroups and so forth. This method also allowed for substitution of missing criteria in their application. Both the list format and the classification tree have been widely used as the case definition for RA in observational research and provide entry criteria for the majority of clinical trials in RA since their publication. However, the mean disease duration in the development cohort was 8 years (30); perhaps in part as a result of this, from early on these criteria were criticised for their lack of sensitivity early in the disease course (76;80;81). As a result of this, patients with early arthritis were much less likely to be eligible for clinical trials and research, leading to a scarcity of knowledge regarding the potential efficacy of new drugs at this stage in the disease course. This has become increasingly important over the last 15 years, with mounting evidence that early aggressive treatment, including the use of novel therapies, can prevent the very erosive, disabling joint disease typified by the 1987 criteria (82-84).

ARA 1987 revised criteria for the classification of RA

Criterion	Definition
A patient is classified as RA if 4/7 criteria are satisfied. Criteria 1-4 must have been present for ≥6 weeks	
1. Morning stiffness	Morning stiffness in and around the joints, lasting at least an hour before maximal improvement
2. Arthritis of ≥3 joint areas	≥3 joint areas simultaneously have had synovitis observed ay a physician
3. Arthritis of hand joints	At least 1 area swollen in a wrist, MCP or PIP joint
4. Symmetric arthritis	Simultaneous involvement of the same joint areas on both sides of the body
5. Rheumatoid nodules	Subcutaneous nodules, over bony prominences, extensor surfaces or juxta-articular regions
6. Serum rheumatoid factor	Positive RF
7. Radiographic changes	Radiographic changes typical of RA in posteroanterior hand and wrist radiographs

Figure 1.5.2 ARA 1987 revised criteria for the classification of RA, adapted from (30)

1.5.3 2010 American College of Rheumatology/European League against Rheumatism classification criteria for RA

The most recent set of classification criteria, developed in 2010, endeavour to address these problems (31). They were developed through a joint global initiative of the ARA, now known as the American College of Rheumatology (ACR), and the European League against Rheumatism (EULAR) and aim to identify, at the time of their first presentation to a rheumatologist, those patients with early inflammatory arthritis who will go on to develop persistent, erosive and potentially disabling RA. Three phases were undertaken and are shown in figure 1.5.3.

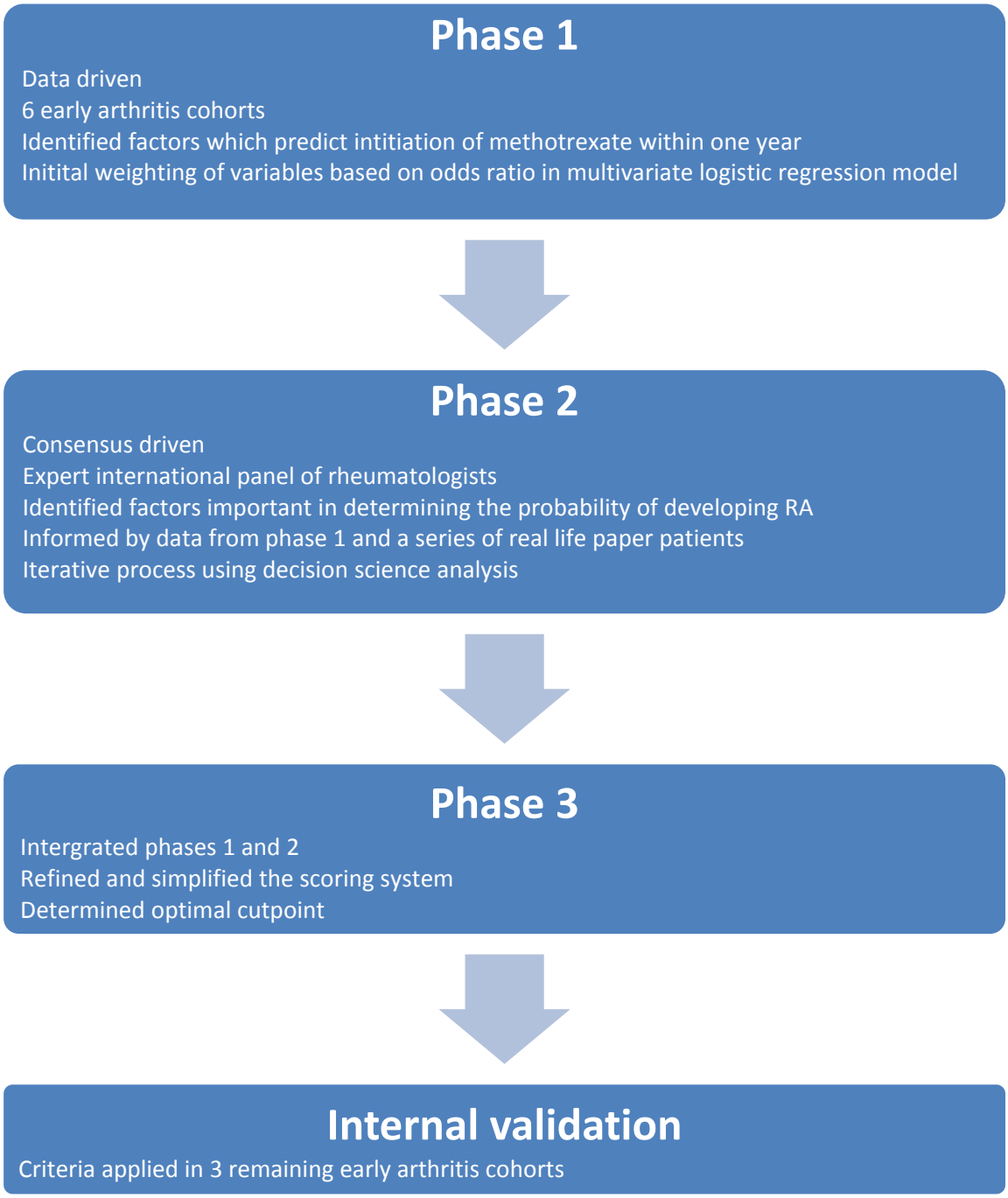


Figure 1.5.3 Phases of development of the 2010 criteria

The first comprised of a data driven analysis of 3115 patients in 9 different early arthritis cohorts from across the world. Six of these cohorts were pooled and used to identify which baseline clinical features and investigations would lead to a patient with synovitis to be commenced on methotrexate (85). Initiation of methotrexate therapy was used throughout the process as a proxy gold standard for the diagnosis of RA, as it should represent those patients the treating clinicians thought

would have a persistent, erosive arthropathy requiring DMARDs. The independent contribution of each variable was analysed using univariate and multivariate logistic regression as well as principal component analysis. Once a set of variables and their weighting had been identified, the second phase was then a consensus driven process using real-life case studies and an international panel of 24 expert rheumatologists (86). A 2 day workshop was held and the domains and factors that were important in the probability of developing RA were discussed, including any other potential important parameters which had not been identified in the data driven phase 1. A decision science software program was then used to determine the relative weights of these domains and allowed the calculation of an individual's likelihood of developing RA. The final phase involved refining the scoring system and determining the optimal cut point (31). The refined final score with weightings gave a potential total score out of 10. To ascertain when, on this scale, a patient should be classified as RA, the case scenarios from phase 2 were ranked according to the new scoring system and each member of the expert panel was asked to identify at which point the cases changed from probable to definite RA. Additionally, the new criteria were then applied to the remaining 3 early arthritis cohorts not originally included in phase 1, and sensitivity and specificity calculated. Receiver operating characteristic (ROC) curves were drawn by plotting sensitivity against 1-specificity. These curves allow selection of the cut point with the highest combination of sensitivity and specificity. Both methods identified the optimal cut point as ≥6/10. The final criteria set are shown in figure 1.5.4.

2010 ACR/EULAR Classification Criteria for RA

Target population: Patients who (i) have at least one joint with clinical synovitis, and (ii) with the synovitis not better explained by another disease	**Score**
Joint involvement (tender/swollen)*	
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
Serology†	
Negative RF&ACPA	0
Low-positive RF/low positive ACPA	2
High positive RF/high-positive ACPA	3
Acute-phase reactants	
Normal C-reactive protein (CRP) & erythrocyte sedimentation rate (ESR)	0
Abnormal CRP/ESR	1
Duration of symptoms	
<6 weeks	0
≥6 weeks	1
Add score of categories A-D: ≥6/10 = definite RA	

***** Large joints were defined as shoulders, elbows, hips, knees and ankles. Small joints are defined as metacarpophalangeal joints, proximal interphalangeal joints, second through fifth metatarsophalangeal joints, thumb interphalangeal joints and wrists. Distal interphalangeal joints, first carpometacarpal joints and first metatarsophalangeal joints were excluded

†RF/ACPA results were classified as negative: defined as ≤ upper limit of normal (ULN) for the laboratory and assay; low positive: > ULN but ≤ 3 times ULN; and high positive: > 3 times ULN

Figure 1.5.4 2010 ACR/EULAR classification criteria for RA, adapted from (31)

The methodology undertaken to develop the 2010 criteria is broadly in keeping with recommendations for consensus methodology (87;88). Consensus methodology is required when guidance is developed in areas where randomised controlled trial

data are not available and expert opinion must be relied upon. The development of classification criteria for RA is a clear example of this. The development phases described above encompass both data driven and consensus methodology, which is commendable. However, it is interesting to note that no systematic literature review was undertaken at the outset, which would be the usual starting point for any consensus process. Although the participants of the expert panel are likely to have been very familiar with relevant literature on the topic, it might have been useful to have that literature presented to the panel systematically before the consensus meeting. In addition, in any face-to-face group discussion, certain members of the group may overshadow others, and their opinions may therefore be overrepresented in the conclusions. This potential problem was minimised by the involvement of an experienced moderator who facilitated the meeting.

1.5.4 Erosive disease in the context of the 2010 ACR/EULAR classification of RA

At the time of their publication, the criteria included an additional item, which stated that a patient could be automatically considered as having RA if there were radiographs available with 'typical RA erosions'. However they explicitly stated that 'a typical RA erosion' had yet to be defined and therefore a further body of work was embarked upon to provide a clear definition of erosive disease. This was undertaken in two of the cohorts used in the development of the criteria, the Leiden Early Arthritis Clinic (EAC) and the Études et Suivi des Polyarthrites Indifférenciées Récentes (ESPOIR) cohort. As with the criteria, the first phase was data driven and looked at the number and site of erosive joints that were associated with the initiation of methotrexate or any DMARD in the first year of disease (89). This generated a number of potential erosive joint count cut-offs with moderate to high specificity (60-95%) that could be taken forward to the second phase. The second phase was consensus based, through face-to-face meetings, online voting and teleconference of the EULAR task force. The results of the first phase were discussed and it was decided to aim for high specificity and focus on patients who did not already fulfil the 2010 criteria (those who scored <6/10). The final unanimous vote decided on the definition of ≥3 erosive joints at any of the following sites: proximal interphalangeal (PIP) joints, metacarpophalangeal (MCP) joints, wrist (counted as one joint) and metatarsophalangeal (MTP) joints on radiographs of both hands and feet. Bilateral affected joints are counted as two joints and three joints in the same joint group (such as MCP) also fulfil the definition (90).

1.6 Validity

After development of a criteria set, validation is crucial. The validity of a measurement (such as a criteria set) can be defined as its ability to measure what it is intended to measure (91), in this instance the ability of the classification criteria to identify which patients have RA. There are a number of different types of validity, such as internal and external validity, content validity, criterion validity and construct validity.

1.6.1 Internal and external validity

Internal validity assesses the validity of a measurement within the cohort in which it has been developed; it measures whether the results of a study reflect the true situation in the study sample. In the case of the 2010 criteria this was tested by dividing the cohorts involved into the 6 used for the data driven analysis part of the criteria development and the 3 other cohorts. This approach of splitting data into a 'training' dataset and a 'test' dataset is commonly used in internal validation of prognostic models (91). However, there were specific inclusion and exclusion criteria defining which patients, from any of the 9 early arthritis cohorts involved, were included in the development sample at all, either as part of the training or test dataset. In order to understand how the criteria perform in wider populations, external validation is required. External validation reflects the generalisability of a measure in the populations in which they were intended to be used, which are likely to be different from the development sample. Any studies validating the criteria since their publication fall into this category (92-101) and these are detailed in the systematic review below. Given the heterogeneity of RA and early inflammatory arthritis, external validation of the 2010 criteria is of particular importance.

1.6.2 Content, criterion and construct validity

Content validity measures whether the criteria include all of the concepts and domains contained within the underlying construct of RA. It is a subjective measure, which was addressed in the development of the criteria in the selection of potential variables to include in the data driven model of phase 1; as well as in consensus based, decision analysis approach of phase 2. However the consensus based approach could only reflect the opinions of the 24 members of the expert panel, and as the 2010 criteria become more widely utilized throughout the field of

rheumatology and rheumatological research, their content validity will remain a topic for debate. A particular example of this is the concept of morning joint stiffness, which was considered as a potential variable for inclusion within the criteria despite showing no association with initiation of methotrexate in the data-driven analysis. It was deliberated by the expert panel and eventually decided it should not be included in the final criteria set (102). Nevertheless it may remain a factor that clinicians and others may consider important when classifying a patient as having RA.

Criterion validity compares the test or criteria to a pre-defined gold standard. This type of validity is very important, but is only useful if such a pre-defined gold standard already exists. In the classification of RA, there is no gold standard and external validity must be assessed by another method. Construct validity describes the ability of the criteria to accurately identify the underlying construct. A simple test of construct validity is the 'face' validity of the criteria, i.e. do they appear to describe what is commonly recognised as the disease RA, without any statistical analysis required. Construct validity can also be tested in part by assessing the ability of a criteria set to discriminate between two entities (such as those patients with and without RA), and it can be further tested by the evaluating the relationship between the criteria and defined outcomes associated with the underlying construct (such as persistent or erosive arthritis). These concepts have been studied in a number of external validation studies since the publication of the 2010 criteria which are reviewed below.

1.7 Classification criteria in the epidemiology of RA

As well as defining inclusion criteria for clinical trials, one of the other important uses of classification criteria is to allow us to describe important epidemiological properties of a disease. For example, rates of disease occurrence, as measured by incidence and prevalence, can only be accurately obtained where there is a clear definition of the disease. As a result, if classification criteria are updated, or new criteria are developed, it is important that new estimates of disease occurrence are obtained, and comparisons made to estimates using the previous criteria set. In addition key epidemiological trends in important disease outcomes, such as mortality, may change as a result of new classification criteria and require re-estimation.

1.7.1 Incidence and prevalence

Disease occurrence is defined epidemiologically by its incidence and prevalence. The incidence of a disease is the number of new diagnoses occurring within a specific population (for example a country) in a set period of time; most commonly this is a year. The prevalence of a disease is the total number of patients within a specific population with that disease at any one point in time (point prevalence) or within a specific period (period prevalence). New classification criteria will necessitate re-calculation of these measures in RA.

1.7.2 Mortality trends

Classification criteria are also essential to look at epidemiological trends over time, as a consistent case definition is needed to make comparisons. As survival is decreased in patients with RA (103), trends in mortality rates are of particular interest. In recent years, population mortality rates have decreased significantly (104). In parallel, there has been a drive to improved disease control with aggressive treatment (105), which has led to improvements in long term outcomes such as radiographic progression (106). It could therefore be postulated that mortality rates in RA may also have improved.

1.8 Systematic review of observational studies investigating the 2010 ACR/EULAR classification criteria for RA

1.8.1 Aims and search strategy

The aim of this systematic review was to identify, summarise and appraise all full text publications validating or analysing the properties of the 2010 ACR/EULAR classification criteria for RA. The intention was to include specifically validation studies which could be collated and compared, but also to collect and discuss any other studies exploring the criteria, in order to obtain as much information as possible. The words 'classification criteria' and 'rheumatoid arthritis', were searched for as MeSH terms and keywords in titles, abstracts and whole texts, published after August 2010 (the date the criteria were published) in the Medline and Embase databases. The complete search strategies for each database are shown in appendix 1. The search was last updated on 31st December 2014. Titles and abstracts were reviewed and full texts retrieved if they were relevant to the review question. Conference proceedings, case reports and case series were not included, neither were randomised controlled trials (RCTs). The articles describing the development of the criteria and the definition of erosive disease in the criteria were excluded, as were any publications arising from the work presented in this thesis. The bibliographies of selected articles were searched for other relevant publications not identified by the main search strategy. Full texts were only included in the main review if they reported performance characteristics of the criteria. Studies investigating other aspects of the criteria, along with review articles and editorials, were not included in the final selection but were used to inform and enhance the search and discussion of the results. The selection of articles is shown in figure 1.8.1. After full text review, 25 studies remained.

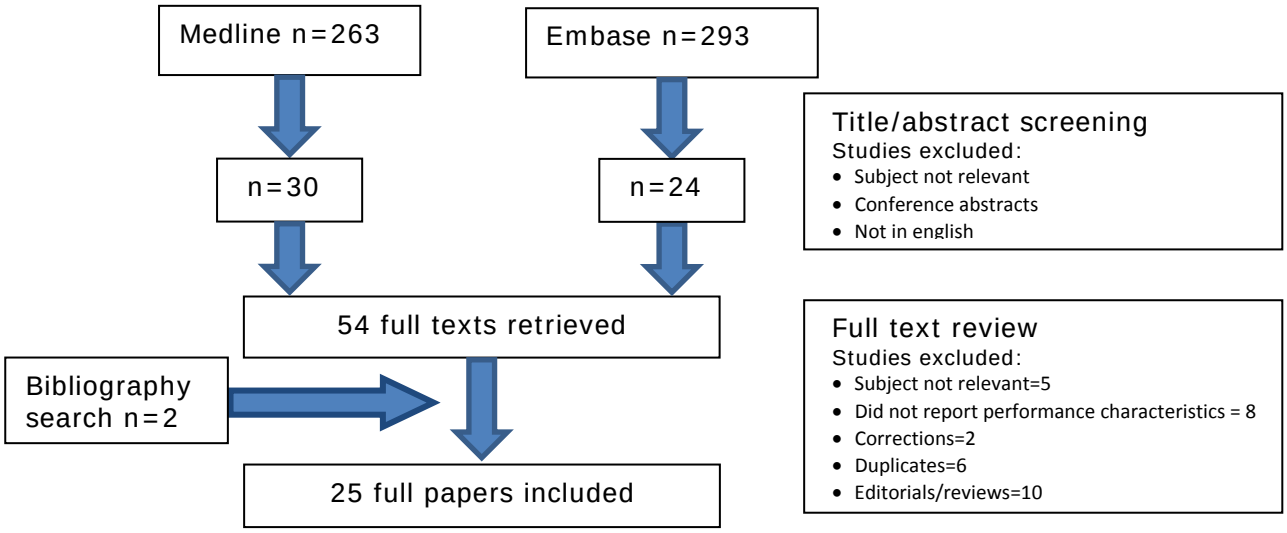


Figure 1.8.1 Flow diagram of literature search

In the included studies the performance of the criteria is explored in relation to a number of different proxy gold standards: initiation of methotrexate/DMARDs, physician diagnosis of RA, and persistent/erosive disease. An explanation of how performance characteristics are calculated and interpreted is given below.

1.8.2 Performance characteristics

A number of different measures are reported in the studies which explore the performance characteristics of the criteria. All the studies included report the sensitivity, which in this context is the number of patients who fulfil the criteria and meet the gold standard as a proportion of all those who meet the gold standard; and specificity, which is the number of patients who do not fulfil the criteria and also do not meet the gold standard as a proportion of all those who do not meet the gold standard. The positive and negative predictive values (PPV and NPV) are, respectively, the proportion of patients meeting the gold standard and fulfilling the criteria out of all those who fulfil the criteria and conversely, the proportion of patients not meeting the gold standard and not fulfilling the criteria out of all those who do not fulfil the criteria (see Figure 1.8.2). The positive predictive value therefore represents the likelihood of a patient having the disease if they fulfil the criteria. Other measures frequently reported are the area under the curve (AUC) and likelihood ratios. The AUC is the area under a receiver operating characteristic (ROC) curve, which plots sensitivity on the x-axis against 1-specificity on the y-axis; therefore to some extent it represents a composite of sensitivity and

specificity. Alternatively, likelihood ratios are the ratio of sensitivity and specificity; thus the positive likelihood ratio (LR+) is the probability of a patient fulfilling the criteria if they meet the gold standard divided by the probability of them fulfilling the criteria if they do not meet the gold standard. As they are ratios, they are not dependent on the total prevalence of gold standard positivity in the cohort, unlike PPV and NPV. An optimal diagnostic test would have a LR+ >10 and a LR- of <0.1. All these measure thus provide different information about the overall performance of the classification criteria. They can all be calculated from a simple 2x2 contingency table; the formulae for these calculations are shown below (figure 1.8.2).

Criteria	Gold standard	
	+ve	-ve
+ve	a	B
-ve	c	D

Sensitivity = a/a+c **PPV** = a/a+b

Specificity = d/b+d **NPV** = d/c+d

LR+ = (a/a+c)/(1-(d/b+d))

LR- = (1-(a/a+c)/(d/b+d))

Figure 1.8.2 Formulae for calculation of performance characteristics

1.8.3 Results

The inclusion and exclusion criteria of the cohorts in the selected studies are shown in table 1.8.1, along with baseline demographics and disease specific variables. Table 1.8.2 shows the same details for other relevant studies which did not specifically report performance characteristics. The results of the studies from table 1.8.1 are then discussed grouped by the gold standard used in the analysis. It should be noted that some studies assessed the criteria using more than one outcome measure.

Table 1.8.1 Characteristics of cohorts in validation studies

Authors	Year study first recruited patients	Cohort size *n*	Inclusion criteria	Age *(years)*	Female *n(%)*	Disease duration *(months)*	Follow up
Alves (92) Netherlands	2004	513	• ≥16yrs • ≥1 swollen joint • Or ≥2 painful joints • PLUS ≥2 of o I hr morning stiffness o Unable to clench fist o Pain on shaking hands o Paraesthesia in fingers o Unable to wear rings o Family history RA o Unexplained fatigue • Atraumatic • <1 year symptom duration	50(14)*	374(73)	3.5(0.03-12)‡	0,6,12 months
Berglin (107) Sweden	2004	313	• ≥1 swollen joint • No alternative diagnosis • <1 year symptom duration • ≥1 year follow up	60(46-70)†	203(65)	4(3-7)†	retrospective, 13(12-15) months†
Biliavska (108) International	2004	303	• ≥1 swollen joint • <16 weeks symptom duration • DMARD naïve	48(16)*	109(76)	57(30-88) days†	2, 12 & 52 weeks
Britsemmer (93) Netherlands	2000	455 (175 had radiographic progression data)	• ≥18yrs • ≥2 swollen joints • <2 years symptom duration • DMARD naïve	52(13)*	314(69)	5.6(5.5)*	0, 1 & 3 years
Cader (94) UK	Not stated	205	• ≥1 swollen joint • <3 months symptom duration • ≥18 months follow up	49(35-64)†	118(58)	1.4(0.83-2)†	0, 1, 2, 3, 6, 12, 18 months

***mean (+/-standard deviation)**
†median (interquartile range)
‡ median (range)

Authors	Year study first recruited patients	Cohort size *n*	Inclusion criteria	Age *(years)*	Female *n(%)*	Disease duration *(months)*	Follow up
Cornec (95) France	1995	164	• ≥16yrs • ≥1 swollen joint • No prior joint diagnosis • <1 year symptom duration	50(5)*	117(71)	10.4(3.7)*	0,6,12,18,24 months & 10/12 years
de Hair (109) Netherlands	2002	301	• <1 year symptom duration	50(18-88)‡	189(63)	4(0-12)‡	0 & 2 years
Kaarela (96) Finland	1973	121 "true RA" 95 controls	• ≥16yrs • ≥1 swollen joint • <6 months symptom duration	46(13)*	81(67)	Not stated	0,1,3,8,15,20 years
Kaneko (97) Japan	2009	313	• Joint symptoms (arthralgia/joint swelling/morning stiffness) • DMARD naïve	54(14-86)‡	247(79)	4.2(0.2-243)‡	Cross sectional, retrospective
Kasturi (110) USA	1976 Nurses' Health Study I (NHSI) 1989 Nurses' Health Study II (NHSII)	128 new self-reported RA from 121 700 participants (NHSI) and 116 608 participants (NHSII)	• Female nurses • Aged 30-55 at inception(NHSI) • Aged 25-42 at inception (NHSII)	NHSI: 71(7)* NHSII: 54(5)*	NHSI: 39(100) NHSII: 89(100)	Not stated, incident self-reporting	Cross sectional, retrospective
Kawashiri (111) Japan	2010	69	• Arthritis • Alternative diagnosis excluded • <1 year symptom duration	54(17)*	54(78)	4(3)*	3, 6, 9, 12 months
Kennish (112) USA	2010	126	• Any joint symptoms	48 (15)	87(78)	5.3 years	cross sectional, retrospective

***mean (+/-standard deviation)**
†median (interquartile range)
‡ median (range)

Authors	Year study first recruited patients	Cohort size *n*	Inclusion criteria	Age *(years)*	Female *n(%)*	Disease duration *(months)*	Follow up
Kim (113) South Korea	2009	75	• First visited outpatient rheumatology 2009-2010 • Joint symptoms • Bone scintigraphy available • Arthritis 1 joint	45(13)*	124(80)	29(47)*	Retrospective
Makinen (114) Finland	1997	221 with all data available from cohort of 377	• Recent onset synovitis • Not better explained by another disease • In follow up for ≥2 years • Responded to invitation for 10 year follow up visit	53(14)*	154(70)	6(3-13)†	10 years
Mourao (99) Portugal	2005	37	• ≥4 swollen joints • <6 weeks symptom duration • DMARD & steroid naïve	48(18)*	25(68)	Not stated	33(11) months *
Nakagomi (115) Japan	2010	109	• Musculoskeletal symptoms • ≤3 years symptom duration • Referred to outpatient immunology clinic • No alternate diagnosis	52(15)*	85(78)	24(12-40) †	1 year
Neiuwenhuis (116) Netherlands	2010	205	• Arthritis confirmed by rheumatologist • <2 years symptom duration • Did not satisfy the 1987 criteria at presentation	55(15)*	125(61)	11(5-25)‡ weeks	1 year

***mean (standard deviation)**
†median (interquartile range)
‡ median (range)

Authors	Year study first recruited patients	Cohort size *n*	Inclusion criteria	Age *(years)*	Female *n(%)*	Disease duration *(months)*	Follow up
Raja (117) New Zealand	2004	79	• Joint pain and stiffness • ≥1 swollen joint • <12 months symptom duration • In follow up for ≥1 year	58(23-84)‡	53(67)	3.6(0.5-12)‡	Ever 3-6 months for the first year, then annually
Ravindran (118) India	2011	134	• Age 18-75 • ≥1 swollen joint • <1 year symptom duration • No alternate diagnosis	not stated	not stated	not stated	1 year
Reneses (119) Spain	2002	201	• 16yrs • ≥2 swollen joints • ≥4 weeks <12 months symptom duration • No alternate diagnosis • DMARD and steroid naïve	51(17)*	144(72)	6.3(3.8)*	1 year
Tamai (120) Japan	2001	166	• Arthritis confirmed by rheumatologist • Not classifiable according to ACR criteria within 2 weeks after being included in the study	not stated	not stated	not stated	1 year
Tamas (121) Romania	2009	64	• Age 18-75 • ≥1 swollen joint • <1 year symptom duration • No alternative diagnosis	45	f/m 3:1	3*	not stated

***mean (standard deviation)**
†median (interquartile range)
‡ median (range)

Authors	Year study first recruited patients	Cohort size *n*	Inclusion criteria	Age *(years)*	Female *n(%)*	Disease duration *(months)*	Follow up
Van der Linden (100) Netherlands	1993	2258	• Arthritis confirmed by rheumatologist • <2 years symptom duration	52(17)*	1340(59)	• <6 weeks n=436(21%) • ≥6 weeks n=1602(79%)	0 & 1 year
Varache (101) France	1995	270	• ≥18yrs • ≥1 swollen joint • No prior joint diagnosis • <1 year symptom duration	Not stated	Not stated	Not stated	30 months∞
Zhao (122) China	2009	404	• ≥1 swollen joint • Outpatient clinic attendance • In follow up for ≥1 year	50*	269(67)	24 (0-600)‡	1 year

***mean (standard deviation)**
†median (interquartile range)
‡ median (range)

Table 1.8.2 Cohort characteristics of studies investigating other properties of the 2010 criteria

Authors	Year study first recruited patients	Cohort size *n*	Inclusion criteria	Age *(years)*	Female *n(%)*	Disease duration *(months)*	Follow up
Burgers (123)	1993	1502	• Arthritis confirmed by rheumatologist • <2 years symptom duration • >1 year follow up • Not taking part in RCTs	57(16)	68	20(11-37) weeks	Annually for 7 or 10 years
Bykerk (124)	2007	1187	• ≥16yrs • ≥2 swollen joints OR 1 swollen MCP/PIP plus o RF positive o ACPA positive o Morning stiffness ≥45 mins o Response to NSAIDs o Painful MTP squeeze test • ≥6 weeks <12 months symptom duration • No alternative diagnosis	53(15)*	863(73)	6.1(3.2)*	3 monthly for the first year, 6 monthly thereafter
Fautrel (125)	2002	811	• 18-70 yrs • ≥2 swollen joints • ≥6 weeks <6 months symptom duration • DMARD/steroid treatment <2 weeks duration • Alternative diagnoses exclude	48(13)*	624(77)	Not stated	2 years
Krabben (126)	1993	2472	• Arthritis confirmed by rheumatologist • <2 years symptom duration	Not stated	Not stated	Not stated	Annually for 7 years

***mean (standard deviation)**
†median (interquartile range)
‡ median (range)

Authors	Year study first recruited patients	Cohort size *n*	Inclusion criteria	Age *(years)*	Female *n(%)*	Disease duration *(months)*	Follow up
Jung (127)	2009	102	• ≥18yrs • Physician diagnosis of undifferentiated arthritis • No erosions	47(14)*	73(72)	• <6 weeks n=17(17%) • ≥6 weeks n=85(83%)	20(13) months*
Mueller (128)	1998	592	• RA or undifferentiated arthritis diagnosed by rheumatologist • <367 days symptom duration • DMARD naïve • At least one follow up visit	54(15)	399(67)	177 days*	44(1-178) months‡
Saraux (129)	2002	692	• Age 18-70 • ≥2 swollen joints for ≥6 weeks • ≤6 months symptom duration • No prior DMARDs • No prior steroids unless ≤2 weeks at mean dose ≤20mg/day	48(12)*	528(76)	75 days*	6 monthly for the first 2 years and annually thereafter
Tamas (130)	2009	100	• Age 18-75 • ≥1 swollen joint • <1 year symptom duration • No alternative diagnosis	44(15)*	f/m 3:1	4(3)*	not stated

***mean (standard deviation)**
†median (interquartile range)
‡ median (range)

Authors	Year study first recruited patients	Cohort size *n*	Inclusion criteria	Age *(years)*	Female *n(%)*	Disease duration *(months)*	Follow up
Van der Linden (131)	Leiden: 1993 Berlin: 2004 NOR-VEAC: 2004	Leiden: 625 Berlin: 154 NOR-VEAC: 193 Total: 972	Leiden: • Arthritis confirmed by rheumatologist • <2 years symptom duration Berlin: • ≥2 swollen joints • > 4 weeks & <12 months symptom duration NOR-VEAC: • ≥1 swollen joint • <16 weeks symptom duration	Leiden: 51(17)* Berlin: 51(15)* NOR-VEAC: 46(15)*	Leiden: 368(59) Berlin: 110(71) NOR-VEAC: 114(59)	Leiden: 5.7(6)* Berlin: 4.6(3.2)* NOR-VEAC: 1.2(1)*	Leiden: annually for 7 years Berlin: 0 & 1 year NOR-VEAC: 0, 3, 6 & 12 months
***mean (standard deviation)** **†median (interquartile range)** **‡ median (range)**							

1.8.4 Initiation of methotrexate/DMARDs

The first approach in validating the classification criteria is to replicate the methodology used in the development of the criteria, i.e. assess the ability of the 2010 criteria to identify those patients who are initiated on methotrexate therapy within one year of symptom onset. This has been done in eight early arthritis cohorts (92-94;100;107;108;119;121). In the internal validation process in the development of the criteria, sensitivity of the 2010 criteria was 87-97% amongst the three cohorts (31). In other early arthritis cohorts, the criteria have not always performed quite as well, as shown in table 1.8.3, with sensitivity ranging from 68-85%. The likely explanation for these findings is the increased heterogeneity amongst patients included in these studies, such as variability in symptom duration at presentation, different minimum numbers of swollen joints, and different follow up schedules. In contrast, the patients included in the internal validation process had to satisfy more rigorous inclusion criteria and as a result were much more homogenous and similar to those patients from whom the variables in the criteria were identified. In most of the studies, specificity was lower than sensitivity, ranging from 50-80%. Given that the new criteria aim to classify patients earlier, it is not altogether surprising that some specificity is lost to gain sensitivity as it is in the very early stages that it is most difficult to differentiate RA from non RA patients.

Table 1.8.3 Performance characteristics of 2010 ACR/EULAR classification criteria for RA - outcome measure: methotrexate use within 1 year

Author	Sensitivity	Specificity	PPV	NPV	LR+	LR-	AUC
Alves (92)	**74%**	**66%**	**76%**	**63%**	-	-	-
Berglin (107)	**84%**	**54%**	-	-	-	-	**0.74**
Biliavska (108)	**84%**	**54%**	**57%**	**82%**	-	-	-
Britsemmer (93)	**85%**	**50%**	**86%**	**49%**	**1.7**	**0.3**	**0.78**
Cader (94)	**68%**	**72%**	**57%**	**81%**	**2.44**	**0.44**	-
Reneses (119)	**80%**	**62%**	**90%**	**43%**	**2.09**	**0.32**	-
Tamas (121)							
Early onset (≤45 years)	**78%**	**75%**	**78%**	**75%**	**3.11**	**0.29**	-
Late onset (>45 years)	**83%**	**80%**	**95%**	**50%**	**4.16**	**0.20**	-
Van der Linden (100)	**84%**	**60%**	-	-	-	-	**0.72**

If the initiation of methotrexate within 1 year is the gold standard for RA within relatively tightly defined set of patients with early inflammatory arthritis, it seems

that a logical slightly broader gold standard of 'initiation of any DMARD' might be more applicable to the broader range of patients external validation mandates. In particular this may be relevant where historical cohorts, set up before methotrexate was accepted as the anchor DMARD in approximately the year 2000, are used for validation.

Table 1.8.4 Performance characteristics of 2010 ACR/EULAR classification criteria for RA - outcome measure: any DMARD use within 1 year

Author	Sensitivity	Specificity	PPV	NPV	LR+	LR-	AUC
Biliavska (108)	80%	61%	82%	71%	-	-	-
Cader (94)	62%	78%	75%	66%	2.81	0.49	-
Kaneko (97)	74%	71%	93%	27%	2.6	-	-
Raja (117)*	78%	83%	97%	29%	-	-	-
Van der Linden (100)	74%	74%	-	-	-	-	0.74

*reference standard was DMARD use within 2 years

In studies to date the 2010 criteria appeared to be slightly less sensitive to this outcome, compared to methotrexate use, but with the benefit of improved specificity (table 1.8.4). The PPV for this outcome was also in general a little higher. Despite the potential effect of historical prescribing regimes, the study from the Leiden EAC demonstrated lower sensitivity to this gold standard than the initiation of methotrexate (100).

Overall, from these studies, there remain concerns regarding the potential for both over- and under-classification. There did not appear to be a pattern of improved performance in the cohorts which investigated patients presenting earlier; in fact the study by Cader et al which captured patients with the shortest symptom duration of less than 12 weeks (94) demonstrated the lowest sensitivity for either gold standard. This suggests that the 2010 criteria are not as yet able to accurately identify RA patients during 'the window of opportunity' (69;70), the proposed period of time immediately after onset of symptoms when inflammation in the joints which has the potential to become persistent and destructive, may be switched off by appropriate immunosuppressive therapy. This re-emphasises that the classification criteria should not be used for diagnostic purposes. The best overall performance characteristics were seen in the study by Tamas et al (121), which recruited patients from a tertiary referral centre where it is likely that some pre-selection of the patients had already occurred.

There are some limitations to these studies. The cohorts have different characteristics, which limits direct comparisons between them. The majority of studies recruited patients from rheumatology outpatient departments and therefore may be susceptible to an element of pre-selection from referrals from the primary care practitioners (94). It should be noted that the studies by Alves et al (92) and van der Linden et al (100) involved cohorts which had contributed some patients to the development of the 2010 criteria. This could have created circularity within the results; however sensitivity analyses were performed excluding those patients and the authors reported no significant alteration in the results. The studies by Raja et al (117) and Tamas et al involved only small numbers of patients, which in the latter study was further reduced by stratifying patients by age (130), it is therefore difficult to draw firm conclusions from these studies.

1.8.5 Physician diagnosis of RA

The use of physician diagnosis as a gold standard is plagued by inherent circularity with the previous criteria set, which, by virtue of their ubiquity in RA literature over the past 20 years, are so ingrained within the physician psyche they are likely to contribute significantly to the diagnostic process. It is because of this circularity, as well as inevitable heterogeneity within this gold standard definition, that physician diagnosis was not used as the gold standard in phase 1 of the 2010 criteria development. Nevertheless amongst the wider rheumatology community it is clearly an important outcome measure.

Table 1.8.5 Performance characteristics of 2010 ACR/EULAR classification criteria for RA - outcome measure: physician diagnosis RA

Author	Sensitivity	Specificity	PPV	NPV	LR+	LR-	AUC
Berglin (107)	91%	65%	-	-	-	-	0.84
Biliavska (108)	85%	64%	74%	78%	-	-	-
Britsemmer (93)	90%	48%	79%	69%	1.73	0.21	0.8
Cornec (95)	58%	89%	74%	81%	-	-	-
Mourao (99)	71%	56%	68%	60%	-	-	-
Kaarela (96)	79%	96%	-	-	-	-	-
Kasturi (110)	79%	87%	95%	55%	-	-	-
Kennish (112)	97%	55%	44%	98%	-	-	-
Ravindran (118)	97%	93%	99%	76%	13.5	0.04	-
Varache (101)	51%	82%	77%	58%	-	-	0.78

In this context, the performance of the 2010 criteria has varied quite widely (table 1.8.5); again this is likely to be because what constitutes RA in the mind of one physician may be markedly different from that of another physician. It is that lack of a specific definition of 'physician diagnosis' that also limits all of these studies. Some of the variation seen may be explained by the setting of the cohorts. For example, the study by Ravindran et al from India reported high performance characteristic in all measures, however it was a small cohort recruited from a tertiary referral centre so there will already have been some selection prior to patients being eligible for recruitment (118). It should also be noted that two of these studies (Cornec et al and Varache et al) reported on the same cohort but used physician diagnosis at different follow up times as their gold standard; Varache et al reported on physician diagnosis after 2 years, whereas Cornec et al reported on physician diagnosis after 10 years. They included different total numbers of patients in each study, so it is difficult to ascertain whether these were completely different subgroups from the same cohort or if some of the patients were included in both studies. In addition, Varache et al (101) reported very few baseline characteristics, with no information on the age and gender distribution of the study sample or symptom duration. Without basic descriptive information it is difficult to interpret the results and how they might apply to other patients. It is interesting that these two studies reported the lowest sensitivities in any of the validation studies across all the different gold standards. If these two studies are excluded the reported sensitivities fall more into line with those in validation studies against other outcome measures, ranging from 70-97%. Specificity however remains extremely variable ranging from 48-96%. Three of the studies partly addressed the inherent circularity of this outcome measure by also investigating the performance of the criteria against other gold standards (93;107;108).

Kennish et al took a slightly different approach, applying the criteria cross-sectionally to a mix of established and newly presenting patients attending a general rheumatology clinic (112). Interestingly, the sensitivity of the 2010 criteria remained high in this context, which might not be expected given the criteria were designed to be applied in early disease. However, their cohort included patients with a range of other rheumatological disorders including SLE and psoriatic arthritis, and as a result there were only 30 patients included who met the gold standard of physician diagnosis of RA. The studies reported by Mourao et al (99) and Kaarela et al(96) were also limited by small sample sizes. In addition, Mourao et al did not report the disease duration of their patients at inclusion in the study; limiting our ability to generalise their results to other patients. Kaarela et al were able to

assess the performance of the 2010 criteria against physician diagnosis after long follow periods, an average of 15 years, which strengthens their results. However, some loss to follow up would be expected over this time, and the authors did not describe how and when this occurred. It is therefore unclear whether attrition bias has influenced the results as it might be expected that patients with RA would be less likely to be lost to follow up than those with a self-limiting inflammatory arthritis.

1.8.6 Persistent disease

The use of persistent arthritis (defined as absence of drug free remission) as the gold standard to assess criteria performance is perhaps the most intuitive of those investigated. It has good face validity and content validity, as persistence of joint inflammation is one of the key concepts of the construct of RA. Three studies have assessed the 2010 criteria against this standard (table 1.8.6).

Table 1.8.6 Performance characteristics of 2010 ACR/EULAR classification criteria for RA - outcome measure: persistent disease

Author	Sensitivity	Specificity	PPV	NPV	LR+	LR-	AUC
Alves (92) (persistent disease at 1 year)	69%	72%	87%	46%	-	-	-
Van der Linden (100) (persistent disease at 5 years)	71%	65%	-	-	-	-	0.65
Zhao (122) (persistent disease at 1 year)	95%	93%	-	-	-	-	0.94

Interestingly the sensitivity and specificity of the 2010 criteria to predict persistent disease, even at different time points, are relatively consistent with findings for other outcome measures. These findings add strength to the overall construct validity of the 2010 criteria. Most of these studies also tested the performance of the criteria against methotrexate/DMARD use and their individual limitations are discussed above. The exception was the study from China by Zhao et al (122), which demonstrated the criteria had very good sensitivity and specificity to identify persistent disease. This may reflect a lower degree of heterogeneity of RA in this ethnic population. In addition, the population in this study had a very wide range of disease duration (up to 50 years) at the time of assessment; this was therefore quite different from the other two studies which looked at early arthritis populations, where the criteria were intended to be applied. An important

limitation of these studies to note is that there is no universally accepted definition of disease persistence.

1.8.7 Erosive disease

Erosive disease also represents a key construct of RA. Three studies to date have validated the 2010 criteria in terms of erosive disease, and in general they have demonstrated high sensitivity and low specificity (93;114). The study by Britsemmer et al consisted of a cohort of DMARD naive patients recruited since 2000. One can therefore assume that the patients were treated according to the modern treatment paradigm of early aggressive therapy and methotrexate as DMARD of first choice, attempting to prevent development of any erosions. Thus the low specificity is to be expected, and the ability of the criteria to identify 88% of patients who do go on to develop erosions after two years and 91% after three years, is a very positive finding. It is these patients in particular that we want to ensure we classify as having RA. It was interesting, however in the study by Makinen et al investigating 10 year erosive disease that the specificity dropped (114); this may simply be a group of patients who do not develop erosions because of aggressive treatment, or may be an indicator of potential over-classification by the 2010 criteria.

Table 1.8.7 Performance characteristics of 2010 ACR/EULAR classification criteria for RA - outcome measure: erosive disease

Author	Sensitivity	Specificity	PPV	NPV	LR+	LR-	AUC
Britsemmer et al (93) (erosive disease at 3 years)	91%	21%	22%	91%	1.16	0.42	0.63
Makinen (114) (erosive disease at 10 years)	87%	44%	68%	72%	1.55	0.29	0.72
Raja (117) (erosive disease at 2 years)	88%	65%	31%	88%	-	-	-

1.8.8 Addition of musculoskeletal imaging

A number of studies have investigated whether including information from musculoskeletal ultrasound (US), magnetic resonance imaging (MRI) or other imaging measures can improve the performance of the 2010 criteria (table 1.8.8).

Nakagomi et al used US detected synovitis instead of clinically detected synovitis of one joint to determine eligibility for application of the 2010 criteria(115); they subsequently used US synovitis in place of the tender and swollen joint counts when applying the criteria. When using US greyscale score of ≥1 they demonstrated sensitivity of 78%, improved from 59% in the clinically derived criteria, with no loss in specificity. However the additional requirement of power doppler (PD) led to a loss in sensitivity and no further gain in specificity. Using an analogous approach, Tamai et al demonstrated that including the presence of MRI bone marrow oedema in patients who did not satisfy the 2010 criteria improved the sensitivity, NPV and accuracy compared to using the criteria alone (120). An electronic communication by Nieuwenhuis et al reported similar methodology to Tamai et al in a small subgroup of patients from the Leiden EAC (116), but found that the increased sensitivity of including bone-marrow oedema came at a significant cost to the specificity. It should be noted that the Leiden EAC study involved only unilateral MRI hand and wrist scans, whereas Tamai et al scanned bilateral hands and wrists. In contrast to these studies, Kim et al used bone scintigraphy, again as a way of additionally identifying joint involvement within the 2010 classification criteria (113). They compared the findings of scintigraphy assisted diagnosis to identify the level of joint involvement, to the use of the 2010 criteria alone. Unlike the MRI and US studies, this method did not demonstrate a marked improvement in performance characteristics.

Kawashiri et al took a different approach. In their study, musculoskeletal US was used after the classification criteria had been applied, in only those patients who had failed to meet the criteria (111). This combination of classification criteria and evidence of PD grade 2 or above on US produced very impressive performance characteristics, although the number of patients in the study was small.

Table 1.8.8 Performance characteristics of 2010 ACR/EULAR classification criteria for RA with the addition of musculoskeletal imaging

Authors & outcome	Sensitivity	Specificity	PPV	NPV	LR+	LR-	AUC
Kawashiri (111) (DMARD use 3 months)	97%	88%	90%	97%	-	-	-
Kim (113) (DMARD use 3 months)	75%	95%	98%	56%	-	-	0.94
Nakagomi (115) (MTX use 1 year)							
GS≥1	78%	79%	-	-	-	-	0.87
GS≥1& PD≥2	56%	94%	-	-	-	-	0.89
Neiuwenhuis (116) (DMARD use 1 year)	83%	45%	57%	75%	1.51	0.37	0.64
Tamai (120) (DMARD use 1 year)	76%	75%	81%	69%	-	-	-

GS: ultrasound grayscale

PD: ultrasound power Doppler

MTX: methotrexate

All of the studies are limited by the smaller numbers of patients typically involved in imaging studies; the results must therefore be interpreted with caution. However, overall it seems the optimal role for musculoskeletal imaging within criteria is not as a replacement for clinical examination, but instead used in patients who do not satisfy the classification criteria based on clinical and serum findings alone. This would also be a more economical approach to the use of imaging which can be costly or require a high degree of operator skill. It also appears that MRI and musculoskeletal US are useful tools but bone scintigraphy may not be.

1.8.9 Comparison with 1987 criteria

A number of studies have compared the performance of the 2010 criteria to that of the 1987 criteria. In general, against all the outcome measures described above, the 2010 criteria appear to have increased sensitivity with decreased specificity (93;94;97;100). In accordance with this, the 2010 criteria appear to be able to classify proportionally more patients at baseline as having RA, whatever outcome measure is used as gold standard (94;109;125;126). De Hair et al used classification according the 1987 criteria after 2 years follow up as the gold standard in their study; they reported sensitivity of 88%, specificity of 76%, with high PPV and NPV (77% and 91% respectively). Of the patients who did not satisfy the 1987 criteria at baseline but did at 2 years, 85% already satisfied the 2010 criteria at baseline. However in the Nurses' Health studies (NHSI and NHSII), the

reverse of this pattern was found, with the 2010 criteria having increased specificity but reduced sensitivity compared to the 1987 criteria (110). This may be due to their study design. NHSI and NHSII are large population cohorts where disease identification occurs through patient self-report, followed by confirmation by clinicians' review of medical records, which may be incomplete. In addition the comparator population are those people who have self-reported RA but this is not confirmed by the clinicians' review, which is different to the other cohorts where the comparator is usually patients with some form of independently confirmed non-RA inflammatory arthritis.

Overall, the majority of evidence suggests that the 2010 criteria do classify more patients earlier on in their disease course than the 1987 criteria (109;124). Notably, however, patients with inflammatory arthritis who meet the 2010 criteria are more likely to achieve spontaneous resolution of symptoms or drug free remission than those who meet the 1987 criteria (94;109;123;126). There is potential therefore, that such patients could be exposed to immunosuppressive therapy they may not require and thus be subjected to unnecessary potential harm. A key criticism of the 1987 criteria was that they restricted access to powerful but beneficial therapies (such as biologics) for patients with early disease. In the new criteria, it is vital therefore to balance the greater availability of, for example, biologic therapy with the known serious risks associated with these treatments (132;133).

There appears to be moderate concordance between the two criteria sets when applied to a single cohort (100;124;125;129). The developers of the 2010 criteria sought to create them independent of the influence of the 1987 criteria, and it has been suggested that the 2010 criteria may be describing a slightly different disease from that defined by the previous criteria set (134). Nevertheless, they are both attempts to describe the construct of RA, so there should be some similarities between the criteria sets. Interestingly, from a biological standpoint, there appears to be little difference in the synovium of patients with RA classified by either criteria set, suggesting any phenotypic differences seen between the criteria sets do not reflect differences in the synovium (98). RA is not confined to the synovium however, and consequences of the disease may be very different elsewhere in the body. It should also be noted that these findings come from a single study, thus may warrant further confirmation.

There is some discordance between the criteria sets, and all of the studies to date have identified subgroups of patients who only satisfy one or other classification criteria set. When the clinical characteristics of these patients are examined, a recurring finding across the studies is that patients who satisfy the 1987 criteria but not the 2010 criteria appear to be much less likely to be positive for RF or ACPA (94;97;110;118;123;125;135), and sensitivity of the 2010 criteria in this subgroup is very low. This suggests that ACPA and RF negative patients in particular may be more susceptible to misclassification as non-RA.

1.8.10 Other analyses

The problem with any of the studies described above is that they are only comparing the performance of the criteria to various proxies that may not fully capture the construct of RA. It should also be acknowledged that any criteria set attempting to classify patients early in the disease course will be imperfect, due to the heterogeneous nature of early inflammatory arthritis. It is useful to consider what other information we can learn about patients with inflammatory arthritis and RA by applying the 2010 criteria. If we consider the total score as a continuous scale from 0-10 of increasing likelihood of RA, rather than treating it as a binary variable using the cut off, then the 2010 criteria can be viewed as a prognostic tool. Thus, the higher a patient scores in the criteria, the more likely it is that in the absence of appropriate treatment, they will develop an erosive and persistent arthritis; that is, the construct known as RA. Comparisons have therefore been made with prediction tools which have been previously validated in early inflammatory arthritis for the purposes of diagnosis (82;136). In this role, as a predictor of RA development, the 2010 criteria perform as well as the previously validated tools (92;93;126), particularly in identifying high risk patients. This may be useful, because the 2010 criteria have been widely published, and could be more readily adopted in practice than several different prediction rules as well as classification criteria. However, when Krabben et al compared categorisation of patients according to the Leiden prediction and the 2010 criteria as a continuous variable, they found that amongst patients categorised as low risk by the Leiden model, there was poor congruence with the 2010 criteria. Over one third of low risk patients according to the prediction model satisfied the 2010 criteria, again highlighting the persistent potential for misclassification.

Two studies have investigated the impact of satisfying the criteria on outcomes over time. Mueller et al compared the evolution of disease activity and

radiographic progression of patients who did and did not satisfy the 2010 criteria, over a median of 3.6 years follow up (128). Interestingly, they found that the cumulative probability of radiographic progression did not differ between criteria positive and negative patients, nevertheless the average change in radiographic damage score at the last follow up was significantly higher in the patients who satisfied the criteria. Disease Activity scores (DAS28) were consistently higher over time in patients who met the criteria, although patients who met the criteria had a greater initial reduction in DAS28 in response to treatment. This apparent better response to treatment might however simply represent regression to the mean. Burgers et al (123), investigated the long term outcomes of patients satisfying the 2010 criteria in comparison to patients satisfying the 1987 criteria. They found that patients satisfying the 2010 criteria were more likely to achieve DMARD free remission and had less radiographic progression than those who satisfied the 1987 criteria.

There have been some investigations into the relative importance of the individual items of the 2010 criteria (95;99;131); and these have highlighted the particular importance of RF and ACPA positivity in predicting outcomes and also improving the test characteristics of the criteria. Van der Linden et al explored this idea further, by examining how much additional sensitivity and specificity was gained by the inclusion of high and low titres of RF and ACPA in the criteria set (131). Interestingly, in 3 early arthritis cohorts from Norway, Germany and Holland, they found no additional value in the inclusion of high positive RF over ACPA positivity, due to a large amount of variation in the sensitivity and specificity of RF amongst the three cohorts. The authors cite the wide variety of techniques available to detect RF as a possible explanation. The ACPA measurements appeared to be more consistent. In their conclusion therefore they suggest that attempting to define high titre RF positivity is complex, and may hamper application of the criteria.

1.8.11 Conclusions

Although the range of cohorts, variety of gold standards and methodological approaches preclude any formal meta-analysis of these studies, a number of themes can be identified. Firstly, amongst early arthritis cohorts, the 2010 criteria do classify more patients at baseline than the 1987 criteria. Secondly, for the majority of studies, whatever the gold standard, sensitivity was usually above 70% and frequently higher. Thirdly, however, it must be acknowledged that the problem

of misclassification (false positives and false negatives) remains a feature of these criteria as much as with the previous criteria set. In particular, specificity does appear to be lower than that of the 1987 criteria, which is logical given the aims of the criteria development were to improve the sensitivity and this often comes at the cost of lower specificity. The use of musculoskeletal imaging as an adjunct to the 2010 criteria to improve their performance is an area with great potential, particularly in patients who do not meet the criteria. However as yet there is insufficient evidence to make solid recommendations. Finally, the subgroups of patient who are classified as RA by only the 2010 or only the 1987 criteria set warrant further detailed investigation to understand their long term prognosis. In particular, lack of autoantibodies in these patients is of interest, as is the broader role of autoantibodies within the criteria. Nevertheless, we may have acquired sufficient knowledge regarding the properties of the criteria to be able to use them within clinical research; perhaps, as suggested by Aletaha et al (31), alongside the 1987 criteria for the foreseeable future.

Although unambiguously defined as classification criteria by the developers, it is widely acknowledged that the 2010 criteria are likely to be used as a guide to diagnosis in clinical practice, and it is valuable to know that they perform similarly to pre-existing and validated diagnostic prediction rules. Further, it has been proposed that any classification criteria can evolve into diagnostic criteria if there is accumulation of sufficient external validation (137). If it is expected that the criteria are adopted as a diagnostic aide, even if not as full diagnostic criteria, it is essential that as thorough external validation as possible is conducted, in order to provide clear information for treating clinicians. For this it is as important to know the long term outcomes of those patients with early arthritis who fail to satisfy the criteria as of those patients who do fulfil the criteria.

1.9 Summary of introduction and literature review

Classification criteria play an essential role in research in RA, as well as providing guidance to clinicians and patients when making diagnostic and prognostic decisions. The 1987 ACR criteria had been demonstrated to no longer be fit for purpose in the modern era of rheumatology, because of their poor performance in classifying patients with early RA. The 2010 criteria are now approaching their 5th anniversary of publication. As shown in the literature review they have demonstrated improved sensitivity in early arthritis compared to the previous criteria set and can predict erosive disease and persistent arthritis. However, the construct validity of the criteria could be further tested by investigating whether they identify patients at risk of long term disease activity, disability and increased mortality. They have been assessed in the observational studies described above and begun to be adopted as entry criteria in clinical trials (138), but have yet to be used to define RA in studies of disease occurrence and epidemiological trends. Further, during their development, the aims of the classification criteria focussed on trying to capture patients when they first present to healthcare. Although the studies by Kennish and Zhao included some patients with disease duration of many years (112;122), these were as part of a mixed cohort of early and established disease; to date there have been no studies which have tested the 2010 criteria exclusively in patients with longstanding inflammatory arthritis. Interestingly, the 1987 criteria were never criticised for their performance in patients with established disease, so the applicability of the 2010 criteria in that group of patients warrants further exploration.

The 2010 criteria place significant emphasis on the presence and level of the two established antibodies RF and ACPA. Two areas of interest involving these autoantibodies have been raised in the literature review. Firstly one study has identified that there may be no additional benefit in assessing the levels of RF to inform prognosis (131); this requires further investigation to look at the effect of antibody levels on other poor prognostic outcomes such as mortality. Secondly it appears that the 2010 criteria potentially miss a group of apparently seronegative patients with poor prognosis. In these patients therefore it may be useful to look at whether novel antibodies such as anti-CarP antibodies are able to provide additional prognostic information.

1.10 Aims and Objectives

1.10.1 Aims

The overarching aims of this thesis are:

- To validate and explore the properties of the 2010 ACR/EULAR classification criteria for RA.
- To characterise the role of autoantibodies in predicting long term outcomes for patients with inflammatory polyarthritis (IP) in the context of the 2010 criteria.

1.10.2 Objectives

The following objectives provide the framework to address these aims:

1. To estimate the incidence of RA as defined by the 2010 criteria in the UK (addressed in chapter 4).
2. To establish whether the 2010 criteria identify, from a cohort of patients with early inflammatory arthritis, those who are at increased risk of early death (addressed in chapter 5).
3. To investigating the role of autoantibody status and levels (as defined in the 2010 criteria) in predicting increased mortality in patients with early inflammatory arthritis (addressed in chapter 5).
4. To utilise the 2010 criteria to describe trends in mortality over time in RA (addressed in chapter 5).
5. To establish whether the 2010 criteria identify a group of patients with early inflammatory arthritis who have a worse long term prognosis in terms of functional disability, disease activity, and radiological outcomes (addressed in chapter 6).
6. To investigate whether patients with IP who are anti-CarP antibody positive have worse prognosis compared to those who are anti-CarP antibody negative in terms of long term functional disability and disease activity (addressed in chapter 6).
7. To discuss the potential challenges in applying the 2010 classification criteria in patients with established RA (addressed in chapter 7).

Chapter 2: Methods

In this chapter methods are described relating to the design of the Norfolk Arthritis Register, as well as the general and specific statistical methods used to address each of the research objectives.

2 Methods

The methodology chapter comprises of three sections. The first describes the methods relating to the Norfolk Arthritis Register, which is the setting for the analyses in this thesis. The second section describes general statistical issues and methodology applicable to all objectives in this thesis. The third section describes the specific statistical methods relating to each of the objectives described above. As this thesis is submitted in alternative format, they will also be described in the manuscripts that make up the subsequent results chapters. However in this chapter, the statistical methods are described in more detail, with explanations of the underlying assumptions and the reasons for selecting each approach to answer the individual research questions.

2.1 Study population

The study population for this thesis is the Norfolk Arthritis Register (NOAR). NOAR is a large primary care inception cohort of patients with IP. Established in 1989, patients are referred to NOAR by their primary care physician or treating rheumatologists if they are adults aged 16 years or more, have two or more swollen joints for at least four weeks. Patients recruited from 1990-2004 could have symptom onset at any time from 1st January 1989 onwards, and for those recruited since 2004 any time form 1st January 2000. There are no other restrictions on inclusion within the study at baseline; throughout follow up patients are excluded if they have any diagnosis (confirmed by a consultant rheumatologist) other than RA, PsA, undifferentiated arthritis or post-infective arthritis. In addition patients are only followed up after 5 years if they met certain criteria (see section 2.2.3 below). The study was approved by the Norwich Local Research Ethics Committee (REC number 2003/075); all prospective participants are provided with written and verbal information and give written consent at baseline, then annually for 5 years and every 5 years thereafter. Originally set up to identify the incidence and outcome of inflammatory arthritis and the subset of these patients with RA, data from the register have led to many publications on prognosis, natural history and time trends in the epidemiology of inflammatory arthritis. A selection of these are summarised in two published review articles (139;140). All patients included in the analyses in this thesis were recruited between 1990 and 2009.

2.1.1 The NOAR cohorts

Recruitment was undertaken in 4 distinct cohorts defined by the year in which patients were registered to NOAR. Follow up also differed between the cohorts (figure 2.1.1), as described in section 2.1.7.

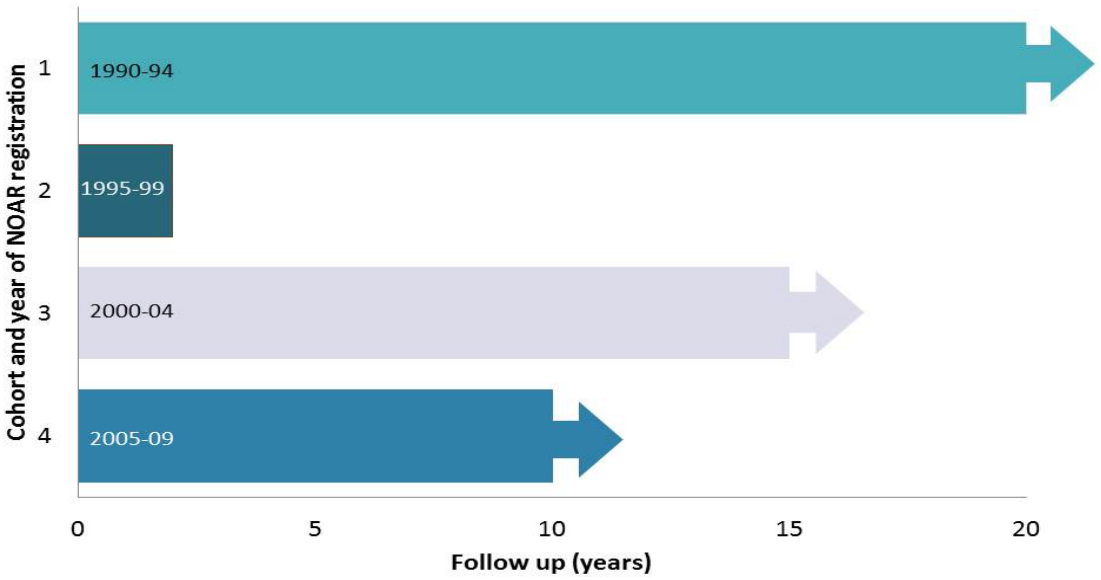


Figure 2.1.1 NOAR cohorts and timing of follow up

2.1.2 Clinical examination and questionnaire

Patients who consented to participate were assessed at baseline, within 2 weeks of notification, by a research nurse who visited them in their home or invited them to a research clinic. The nurse administered a questionnaire, which included details of socio-demographics, symptom onset, lifestyle factors including smoking status, co-morbid disease and any medication including treatment for their joint symptoms. The nurse also performed a 51 tender and swollen joint count and for patients recruited to cohorts 3 and 4, measured height and weight. All patients completed the British version of the Stanford Health Assessment Questionnaire (HAQ) (141). A copy of the baseline NOAR questionnaires (nurse administered and patient completed) is included in appendix 2.

2.1.3 Blood samples

At baseline, Ethylenediaminetetraacetic acid (EDTA) and serum blood samples were taken from all patients who gave consent. DNA was extracted from the EDTA sample and the serum was frozen. The frozen samples were transferred to the laboratory at the Centre for Musculoskeletal Research, University of Manchester for measurement of CRP, RF and ACPA, and storage for any future additional tests. CRP was measured by endpoint immunoturbidimetric agglutination method (Boehringer Mannheim), RF by tube latex dilution test and ACPA IgG antibody detected by ELISA (Axis-Shield Diastat Anti-CCP kit). Further blood samples were taken for RF measurement at any follow up point where the patient who had not yet met the 1987 ACR classification criteria for RA but fulfilled 2 of the 1987 criteria, and therefore with the addition of RF and radiographic data might fulfil the criteria. Repeat serum and EDTA samples were obtained every 5 years from all patients who consented to this. Available stored samples from patients in cohorts 1, 3 and 4 were later sent to Leiden University Medical Center, Leiden, The Netherlands, for testing of anti-carbamylated protein (anti-CarP) antibodies (in-house ELISA based on carbamylated fetal calf serum). A proportion of patients had missing data, particularly ACPA status. Where possible this was supplemented from results in contemporaneous medical records if samples had been taken as part of their routine clinical care.

2.1.4 Radiographs

The intention from the outset of NOAR was for all patients who had RA (as defined by the 1987 criteria) to have radiographs of hands and feet taken. The timing of radiographs differed by cohort. In **cohorts 1 and 2**, it was assumed at baseline there would be no evidence of erosions as patients were recruited early and therefore patients were invited to have radiographs performed at the 1st or 2nd anniversary in (i) all patients who already satisfied the 1987 criteria, or (ii) as described above for RF, where a patient could satisfy the 1987 RA classification criteria if radiographic information were available. All patients in cohort 1 were invited to have radiographs performed at the 5th anniversary. For patients in **cohorts 3 and 4** radiographs of the hands and feet were taken in all patients who consented at baseline or 1st anniversary and additional radiographs were obtained at other follow up assessments where a patient could satisfy the 1987 RA classification criteria if radiographic information were available. Radiographs were

then evaluated using Larsen scoring (142) by two trained independent assessors, with arbitration by a third in the case of disagreement.

2.1.5 Classification

In this thesis, the application of the 1987 and the 2010 classification criteria occurred as detailed in figure 1.5.2 and 1.5.4 (shown in chapter 1), using data collected at baseline assessment. As described in the 2010 criteria, the local laboratory reference ranges dictate the cut-offs for elevated or normal. CRP was considered elevated if >5 mg/l. Values of RF and ACPA are divided into the following groups for scoring: negative: defined as ≤upper limit of normal (ULN) for the laboratory and assay; low positive: >ULN but ≤3 times ULN; and high positive: >3 times ULN. In NOAR, the laboratory ULN was 40 international units (IU) for RF and 5 IU for ACPA. Therefore the cut-offs for the low and high positive groups were >40 and >120 IU respectively for RF; >5 and >15 IU respectively for ACPA. For both criteria sets, if data were missing on any variables, total scores were calculated with the missing variable value taken as zero, and patients said to have met the criteria if they reached the defined cut-offs: ≥6/10 for the 2010 criteria and ≥4/7 for the 1987 criteria. A computerised algorithm was developed for each criteria set once all the relevant data had been entered into the main NOAR dataset.

2.1.6 Deaths

All patients were flagged with the Office for National Statistics (ONS) and the NHS-Information Centre (HSCIC). The HSCIC provided notification of any deaths within NOAR and sent details of the cause of death from the death certificates to the Arthritis Research UK (AR UK) Centre for Epidemiology on a quarterly basis. Cause of death was coded according to the International Classification of Diseases (ICD) version 9 for deaths occurring before 2001 and version 10 thereafter. ICD-9 coded deaths were then mapped to the ICD-10 codes after version 10 was released, to provide consistent classification throughout (143). For any patient who left the UK during follow up, the ONS also provided notification of this with a date of 'embarkation'.

2.1.7 Follow-up

Patients were assessed annually for the first 2 years, then in cohorts 1, 3 and 4 additionally at 3rd, 5th, 7th, 10th, 15th, and 20th anniversaries following baseline assessment. At each follow-up, the research nurse repeated the clinical examination and administered the follow up questionnaire (see appendix 3). The 2010 and 1987 classification criteria were applied cross-sectionally and cumulatively after each assessment. Current disease modifying medication and any changes to medication since last assessment were recorded. This included non-biologic and biologic DMARDs with start and stop dates as well as reasons for stopping. Details of incident comorbidities and changes in smoking status were also recorded. The patient again completed the HAQ. At the 4th, 8th, 12th and 18th anniversary patients who remained under active follow up were sent postal questionnaires which included the HAQ and details of medication changes and new comorbidities. Patients were only followed after 5th year assessment in cohorts 1 and 3 if they met any of the following criteria:

- Satisfy the 1987 ACR criteria cumulatively or cross-sectionally
- Evidence of active arthritis at 2 or more assessments
- History of treatment with DMARDs or steroids

Patients in cohort 2 were also followed beyond the 2nd year assessment if they met the above criteria. Patients in cohort 4 were only followed beyond 2 years if they were participants in the cardiovascular substudy, which had two additional inclusion criteria to NOAR: age 18-65 and onset of symptoms <2 years(144).

2.1.8 Validation of the 2010 ACR/EULAR classification criteria for RA in NOAR

Successful external validation of classification criteria needs to be performed in an appropriate cohort of patients. NOAR has a number of strengths which made it ideal for this undertaking. Firstly, the primary care setting reduced the risk of selection bias to a subset of patients with more severe disease. Complete capture of all new cases was particularly important in estimating incidence rate (objective 1), and significant endeavours were made to ensure all patients presenting with swollen joints were referred to NOAR when it was first established. Secondly, data were collected prospectively on all 4 parameters of the new criteria as well as a wide range of other predictor variables. Thirdly, the new criteria are weighted

significantly towards classifying patients who are autoantibody positive (i.e. RF or ACPA antibody positive). Long term outcomes in autoantibody negative patients are still relatively poorly described, particularly patients who are ACPA negative. Thus the large proportion of ACPA negative patients in NOAR provided an excellent opportunity to study outcomes in this group (objectives 5 and 6). It has linkage to the ONS which provided complete capture of deaths (objectives 2, 3 and 4). Finally, the first cohort of patients in NOAR was recruited during an era when RA was treated much less aggressively and biologic agents were unavailable. These patients have been followed for up to 20 years, and it was therefore possible to examine outcomes such as disease activity, disability and radiographic progression without the confounding of the early intensive treatment regimens routinely used in the modern era (objective 5).

2.1.9 Definition of early inflammatory arthritis

Early inflammatory arthritis (EIA) is defined arbitrarily in the literature, varying widely from <6 months to <4 years symptom duration (96;145). In the NOAR dataset early arthritis has previously been defined as presentation <2 years from symptom onset (76). Therefore, the same definition of less than 2 years from symptom onset has been used in this thesis. This applies to results presented in chapters 5 and 6.

2.2 Statistical methods

All statistical analyses were performed using Stata 12 software package (Stata, College Station, TX, USA). The statistical methods are described briefly in the manuscripts that comprise the results chapters. However, here, there is a discussion of some of the general data handling and statistical issues including confounding and the handling of missing data, which are relevant to all objectives in this thesis.

2.2.1 Power

As a longitudinal study with ongoing recruitment, there was no target sample size within NOAR to power a specific research question. At the outset of this thesis, the register had already recruited over 3500 patients (140). Prior to beginning the

thesis, a power calculation was carried out looking at the comparison between ACPA positive and negative patients as follows: in consecutive samples of 1000 subject there was 80% power to detect an odds ratio (OR) of 1.7 (5% significance) for an outcome with a prevalence of 20% or more. To ensure there was sufficient power, retrospective power calculations were also performed for two of the analyses as examples (objectives 2 and 6), using the *power oneproportion* command in Stata. The results are reported in chapter 3.

2.2.2 Attribution of risk

Different factors may increase or decrease risk of an outcome in different ways. For example exposure to steroids may increase the risk of infection whilst the patient is taking the drug, but this risk may diminish or disappear completely once the patient has stopped taking steroids. On the other hand, having a certain genetic mutation may always increase an individual's risk of developing breast cancer. In the majority of analyses in this thesis, the main 'exposure' or independent variable was whether a patient satisfied the 2010 classification criteria. During the development of the 2010 criteria, the panel recommended that if a patient has satisfied the criteria at some point in time, they are classified as having RA from that point onwards. Therefore an 'ever classified' approach was employed, similar to the genetic mutation and breast cancer example, whereby if a patient ever satisfied the criteria they were always considered to have RA. In chapters 5 and 6, the main independent variable was not the criteria, but the presence or absence of autoantibodies. It has been shown that, in general, in patients with inflammatory arthritis, seroconversion of RF and ACPA from positive to negative rarely occurs (42), and that these antibodies do not appear to vary with disease activity (146;147). Thus, as with the criteria, this exposure was considered to be fixed and patients were considered to remain antibody positive if they had ever tested positive. There are no data yet on whether individuals who are found to be anti-CarP antibody positive remain so over time (personal communication, L Trouw), however there are also no data to suggest the reverse. Therefore for the purpose of the analyses it was assumed that they would behave similarly to the other autoantibodies.

2.2.3 Confounding and co-variate selection

A confounder variable is one which is associated with both the independent and dependent variables in a statistical analysis but is not on the causal pathway

between these two variables. The presence of confounding can result in a biased interpretation of the data. The potential for confounding to occur is one of the key weaknesses of observational data. In a traditional randomised control trial (RCT), or in fact laboratory experiment, conditions in the test group are usually made as identical as possible to those in the control group, or by virtue of randomisation they are equally balanced. However, in observational data there are often systematic differences between the groups being compared. As a result, we must make adjustments for these differences within our models. This is done by including as co-variates any factors which may be associated with both the independent and the dependent variables. In this thesis, the independent variable was usually satisfying the 2010 classification criteria. In chapter 6 and one of the manuscripts in chapter 5, the presence or absence of autoantibodies (RF, ACPA and anti-CarP antibodies) was the independent variable. Due to the nature of these independent variables, the number of covariates included was usually quite small. This is because there were very few factors on which data were collected where there was biological plausibility of a confounding association with both the independent and dependent variable (the outcome under investigation). The majority of factors we might consider including are instead likely to be on the causal pathway between the independent and dependent variables. For example, when investigating the association between antibody status and mortality, we might consider disease severity as a potential confounder as it may be associated with increased mortality and the presence of antibodies. However, antibodies may be present before an individual develops any symptoms of the disease (148) and there are no studies in the literature which demonstrate antibodies developing in patients due to more severe disease, only that those who have them have more severe disease. Therefore the direction of causality is that the antibodies come first and lead to the increased disease severity, as shown in the direct acyclic graph (DAG) (149;150) below (figure 2.2.1). As there is no arrow from disease severity to antibody, we describe disease severity as on the causal pathway or as a path variable, and it is not a true confounder. If such a variable is adjusted for, this would lead to possible overadjustment bias (151).

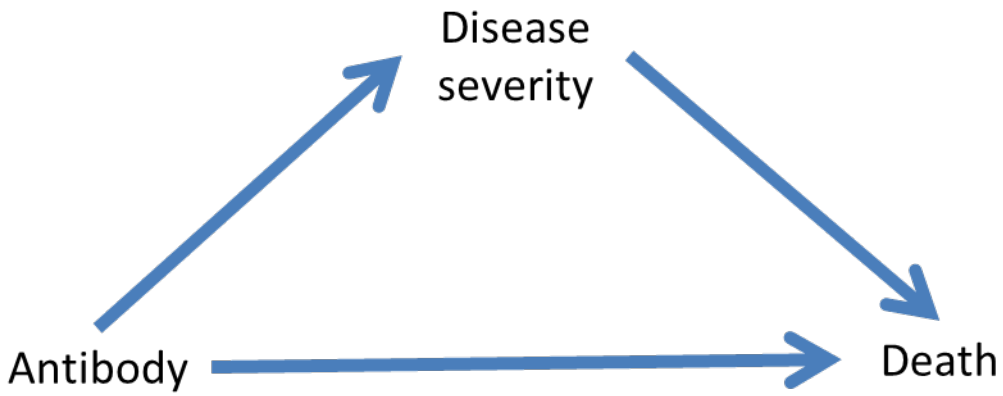


Figure 2.2.1 Example of a direct acyclic graph showing disease severity as a path variable between antibody status and death

Smoking exposure is an example of a true confounder of antibody status and mortality. This is because there is evidence of increased presence of RF and ACPA in smokers (152), and some evidence to suggest that smoking may be a trigger that induces citrullination as described in chapter 1 (47). In addition smoking is well recognised to be associated with increased mortality through a number of other, non-RA related conditions such as lung cancer and cardiovascular disease. In the DAG below which describes this relationship below, in contrast to the first DAG, the arrows of causality travel away from smoking as the confounding variable (figure 2.2.2).

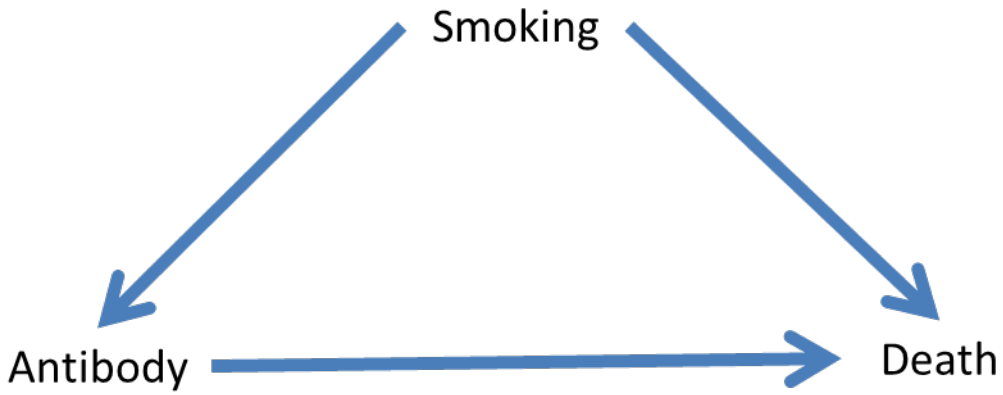


Figure 2.2.2 Example of a direct acyclic graph showing smoking as a confounder variable between antibody status and death

The other problem associated with all observational data, is the issue of unmeasured confounding. A theoretical example might be a hormonal factor occurring during pregnancy or early childhood that causes a persistently mild elevation of ESR (and therefore increasing the chance of satisfying the 2010 criteria), but also independently increases the risk of developing rheumatoid arthritis though triggering the development of autoantibodies. If it is not possible to measure that childhood hormonal factor, this would be an unmeasured confounder and estimates from any regression analysis may be biased by unmeasured confounding. This can be of particular importance when trying to make causal inferences from observed associations. There are statistical methods which can be used to estimate the impact of unmeasured confounding, usually by conducting various sensitivity analyses (153), and in studies investigating the impact of treatment unmeasured
confounding can be adjusted for using propensity scores (154); however these were not employed in this thesis.

2.2.4 Missing data

The phenomenon of missing data is a common difficulty encountered in longitudinal studies and it was important therefore to address this issue and decide for each analysis *a priori* the most appropriate method for handling missing data. Missing data are usually categorized into three main subtypes: missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) (155). MCAR data occur when the probability of a data point being missing is completely unrelated to any of the variables within the dataset. Data are MAR when the probability of a data point being missing is related to one or more of the other variables within the dataset, thus there is some information available in the observed data about the 'missingness'. Data are described as MNAR if the probability of a data point within a variable being missing relates to that variable itself. For example if we were collecting data on whether or not a patient used illegal substances in a questionnaire, it may be that those patients who do use illegal substances are less likely to return the questionnaire, because of their illegal substance use. The presence of missing data is one of the limitations of using a historical cohort to address research questions, such as the validity of the 2010 criteria.

Methods for handling missing data

It is important at the outset of any analysis to decide what type of missing data is present in the dataset and plan the best approach to deal with this. It is useful to compare summary statistics on proposed variables for a regression model between the group of patients with complete data and those with missing data. The simplest approach to deal with the problem of missing data is only to include in a final analysis those patients with complete data on all variables. Complete data is a requirement of any multivariate analysis and participants with incomplete data on any of the covariates will be automatically dropped from the model. This may be acceptable, if the missing data are MCAR and there are therefore no systematic differences between the participants who are included and those who have been excluded. If the data are MAR or MNAR however, a complete case analysis may produce biased estimates. It is also inefficient, as any data collected on incomplete cases are discarded. In addition, the loss of the incomplete cases reduces the sample size which inevitably leads to a less precise estimate (i.e. wider confidence intervals) than would have been possible if all participants could be included.

An alternative to dropping all incomplete cases is to impute missing data. This can be done in a number different ways, all of which have advantages and disadvantages. Examples include simple mean or median imputation, where missing values are replaced by the mean or median of the available data for that variable. However they cannot be employed for categorical data and a consequence of these methods is that the variance of the data is underestimated, which could lead to a biased interpretation of the data. Another example that is commonly used in randomised controlled trials is last observation carried forward, where missing values for an individual are replaced by the last value observed of that variable. The problem with this method is that any analysis of data at the final time point will in fact be an analysis at the last observation, which may have little clinical meaning. An imputation method which is increasingly used for handling missing data is multiple imputation. Multiple imputation is a statistical technique which draws on correlation between the missing variables and those that are available within the dataset.

Multiple imputation

Multiple imputation uses what is known about the variables in which there are missing data, to impute datapoints. Thus, for example there may be missing data on HAQ scores in the dataset. HAQ scores are known to be strongly correlated with age, gender and disease activity, and more complete data are available on these

variables. The data on age, gender and disease activity are used to impute the HAQ scores that are missing, using linear or logistic regression. The method of multiple imputation employed in this thesis was multiple imputation with chained equations (MICE) (156). The chained equation approach imputes the missing data for each variable in turn. Different regression models are used to impute each variable depending on whether the imputed variable is continuous or categorical and the distribution of that variable. Non-normal continuous variables or those with boundaries (such as the HAQ score) need to be normalised prior to running the imputation. Although useful, it should be noted that this technique increases the homogeneity of the dataset. This can be addressed by re-introducing an element of variability prior to performing an analysis, by repeating cycles of imputation. Nevertheless, it is important to consider carefully beforehand whether it is the most appropriate strategy to address missing data.

Missing data in this thesis

In this thesis, four approaches to missing data were taken. Firstly, simple imputation was employed for missing date variables. For any variable that was a date, if the year was missing it was left as missing. If the month was missing it was imputed as the middle month of the year, June. If the day was missing, it was imputed as the 15th. Secondly, for each analysis where there were missing data, characteristics of the patients with and without missing data were examined, and if there were differences, univariate logistic or linear regression analyses were performed to test whether these differences were statistically significant. Thirdly, MICE with 20 imputation cycles was used to produce an imputed dataset in chapter 5, in the analysis addressing objectives 3. A complete case analysis was performed and reported first, and the MICE imputed dataset was then used in a sensitivity analysis to further validate the results of the original analysis. Non- normally distributed variables were normalised using the Stata command *nscore* prior to imputation and converted back after the imputation cycles using the command *invnscore*. Finally, regression models which allow for some missing follow up data (described below in section 2.3.7) were utilised in chapter 6 to address objectives 5 and 6.

2.3 Statistical models

Here, more detail is provided about each of the specific statistical methods used within the results chapters, particularly on the reasons for choosing each of the

models, the strengths and weaknesses of the different models and the assumptions that are made within in them.

2.3.1 Incidence and Poisson regression

The first objective of this thesis was to re-estimate the incidence of RA, given the new definition provided by the 2010 criteria. To calculate the incidence rates, Poisson regression modelling was used. Poisson regression modelling is used for data based on counts, particularly where the outcome is a rare event. It assumes that the occurrence of the disease (or the log transformation of this) changes in a linear way in relation to the exposure variable. In the case of modelling the incidence of RA no exposure variables were included, so this was not important. The Poisson model also assumes that the observations (i.e. new cases of RA) occur independently of one another, and there was no reason to suppose this would not be the case.

2.3.2 Mortality and the 2010 RA classification criteria

Objectives 2, 3 and 4 all propose questions relating to mortality in inflammatory arthritis and are addressed in chapter 5 of this thesis. Objective 2 aimed to test the hypothesis that patients who fulfilled the 2010 ACR/EULAR classification criteria for RA had decreased survival compared to those who did not. This was addressed using Cox proportional hazards model. The Cox proportional hazards model is an adaptation of the Poisson regression model which is used for count data where censorship occurs, that is where an event can only happen once, or the outcome of interest is the first event. Classically it is used in survival analyses to identify predictors of increased or decreased survival, by examining the time-to-event. It models the hazard rate, and in the instance of a mortality analysis this function can be interpreted as the probability if an individual survives to time *t* that they will die the next moment. The hazard ratios produced by the model can therefore be interpreted as the relative risk of death occurring at time *t*. The key assumption in the Cox model is the proportional hazards assumption. This states that the risk of death in the two groups that are being compared are proportional to each other and that this proportion does not change over time. That is, the effects of the predictor variable do not change over time. The *estat phtest* command was run in STATA after each model to test the proportional hazards assumption based on the Schoenfelds residuals. A non-significant result indicates that the proportional hazard assumption has not been violated.

Cox proportional hazards models were also used to investigate the individual components of the classification criteria as predictors of mortality, in order to establish which components were most important in driving this association. Further analyses also looked at how levels and numbers of the antibodies which form part of the 2010 criteria might predict survival (objective 3). In these models it was essential to be particularly selective about which other variables to include as confounders, as many of the factors associated with mortality are in fact more likely to be on the causal pathway rather than a true confounder as explained above.

To calculate changes in mortality rates over time in patients classified as having RA according to the 2010 criteria (objective 4), standardised mortality ratios (SMRs) were used. SMRs compare the observed number of deaths within a pre-specified population (in this instance, patients with EIA) with the expected number of deaths in a sample of the general population with the same age and sex structure followed for the same period of time. As all the patients recruited into NOAR were from the Norwich Health Authority region, we used the local population mortality rates, rather than national mortality rates which might provide a less specific comparison. Data on Norfolk mortality rates were obtained from the Health and Social Care Information centre (HSCIC) (previously National Health Service Information Centre (NHS-IC)) who collect annual mortality data and provided sex grouped rates by 10 year age bands for each calendar year. The NOAR population were also grouped by sex, the same 10 year age bands and calendar years, and the ratio between them calculated. They should be interpreted with caution however, and notably SMRs from different calendar time periods cannot be directly compared as the background referent populations may not be the same. Therefore to investigate potential changes in mortality over time, mortality rate ratios (MRRs) were used. These utilise the Poisson distribution and allow for comparison of mortality rates that had already been standardised to the age, sex and calendar time-specific rates in the general population. The assumption of a linear or log relationship between the independent variable (RA as defined by the 2010 criteria) and mortality rates again seemed reasonable to make. A further relevant assumption made when using Poisson regression in this analysis is that additional covariates, in this case age, gender and disease duration, have a multiplicative effect on the outcome variable.

2.3.3 Associations in longitudinal data: Generalised estimating equations and random effects modelling

In order to investigate predictors of long term outcomes, different statistical approaches can be applied. In the most commonly used approach, standard linear or logistic regression models can study the association between predictor variables and outcomes either cross-sectionally or between a maximum of two time points. For example, in a study with annual follow up over 15 years, we could investigate the association between baseline swollen joint counts and disability measured at the 15 year follow up visit. However, in such a model, any of the data collected between these two time points are discarded. Many disease parameters in chronic conditions such as RA affect patients over the entire course of their disease. Thus an outcome measure, such as HAQ score, recorded at a single point in time may not provide a true reflection of the cumulative burden of disability on the individual throughout that 15 year period. It might be more interesting to know whether swollen joint count at baseline affects a patient's disability throughout their time in the study. However if all of the follow up HAQ scores are included, there are 'repeated measures', and there is likely to be significant correlation of the HAQ scores within individuals. Different statistical methods can be used to look at such question with longitudinal data, and in this thesis, two were used: generalised estimating equations (GEE) (157) and random effects modelling (REM) (158). Both these methods allow inclusion of outcomes measured repeatedly during follow up, but have slightly different strategies for handling the potential for within individual correlation.

GEE is an extension of the generalised linear model (GLM) and compares population averages of the outcome of interest (such as HAQ) in one group compared to another in longitudinal data. It is able to model continuous and categorical variables in a semi-parametric approach and allows the covariance of the within person correlation to be specified (i.e. the way in which the outcome measure is expected to correlate within an individual can be specified in the model).

REM compares the same outcome in an individual participant if they have been exposed to the independent variable, to the outcome in the same participant if they had not. Thus, while GEE produces population averaged estimates, REM yields subject specific estimates. REM is therefore particularly useful where an individual may change their independent variable status, for example stopping smoking.

The main strength of GEE versus REM is that it allows calculation of robust standard errors, and in examples where the question of interest is in population averages in two groups it is the most appropriate model. If subject specific estimates are important REM is more appropriate. In addition, although both GEE and REM are able to include cases with some loss to follow up or other missing follow up data, in GEE it is assumed the missing data is MCAR, whereas REM is able to produce valid inferences with missing data that is both MCAR and MAR (159).

GEE was used in this thesis to address objective 5: do patients who fulfil the 2010 criteria at baseline have worse disease activity, radiographic damage and disability throughout follow up than those who do not and objective 6: do patients who have anti-CarP antibodies have worse disease activity and disability throughout follow up than those who do not. In both the aim was to examine the population average outcomes in each group, therefore GEE was selected as the primary regression model. As described above, the independent variables in these analyses were either criteria status or antibody status and were not expected to change over time. Therefore we could anticipate similar results from both GEE and REM, and REM was used in sensitivity analyses as a way to provide additional robustness to the results.

Chapter 3: Results

NOAR Cohorts

In this chapter the baseline characteristics of the NOAR cohort are presented, as well as other general descriptive data of the cohort including retention within the register over time. These data relate to the whole cohort and therefore inform all of the subsequent results chapters which describe different subsets of the NOAR population.

3 Results: NOAR Cohorts

3.1 Baseline Characteristics

Over 3500 incident cases of inflammatory arthritis have been recruited to NOAR since its inception. Baseline characteristics of the four cohorts with proportions of missing data on any of these characteristics are shown in table 3.1.1. In the whole cohort of 3672 patient, 2410 (66%) were female, median age at symptom onset (IQR) was 55 (43-67) years. The median duration of symptoms (IQR) was 35 (18-83) weeks. A total of 2069 (56%) fulfilled the 2010 ACR/EULAR classification criteria for RA at baseline, 1572 (43%) fulfilled the 1987 ACR criteria.

3.1.1 Differences in baseline characteristics by cohort

There are differences in the characteristics of the patients recruited into the four cohorts (table 3.1.1). As shown below, and demonstrated in a recent analysis (160), baseline disease activity has diminished over the time NOAR has been recruiting patients. Further, there appears to be a reducing trend in the percentage of current smokers and there has also been an increase in the proportion of patients referred by a rheumatologist and a consequent reduction in the proportion referred from their GP. This is shown clearly in table 3.1.1. Perhaps as a consequence of this, there has been a reduction in the proportion of patients recruited in the first 2 years of experiencing symptoms (classified as EIA in this thesis). In addition, the percentage of patients already on DMARDs at the baseline assessment has increased over time. However while the percentage of patients with EIA appears to have plateaued in cohorts 3 and 4, the percentage on DMARDS at baseline continued to increase. This suggests there has also been a change in prescribing practice to earlier initiation of DMARDs.

Table 3.1.1 Baseline demographics and clinical characteristic of NOAR cohorts

	Cohort 1	Missing *n(%)*	Cohort 2	Missing	Cohort 3	Missing	Cohort 4	Missing	Total
n (%)	1098(30)		1093(30)		823(22)		658(18)		3672
Female *n (%)*	715(65)		727(67)		540(66)		428(65)		2410(66)
Median age in years at onset (IQR)*	54(41-67)	-	55(44-67)	-	58(47-70)	-	57(45-68)	-	55(43-67)
Median symptom duration in weeks (IQR)*	22(12-41)	-	28(16-52)	-	29(17-51)	-	27(17-48)	-	35(18-83)
Current smokers *n (%)*	293(27)	1(<0.1)	282(26)	12(1)	191(23)	12(1)	144(22)	3(0.4)	903(25)
Median HAQ (IQR)	0.75(0.25-1.375)	12(1)	0.75(0.25-1.375)	10(1)	1(0.375-1.625)	20(2)	0.875(0.375-1.625)	3(0.4)	0.75(0.25-1.5)
Median DAS28 (IQR)	3.93(2.89-5.01)	221(20)	3.59(2.65-4.68)	245(22)	3.59(2.68-4.46)	166(20)	3.76(2.91-4.69)	165(25)	3.63(2.62-4.68)
RF positive *n (%)*	267(28)	140(13)	278(29)	134(12)	268(37)	90(11)	297(47)	27(4)	1110(34)
ACPA positive *n(%)*	184(22)	280(26)	230(26)	203(19)	147(30)	329(39)	159(33)	171(26)	720(27)
Median CRP (IQR)	5(0-16)	221(20)	8(0-19)	245(22)	10(3-20)	166(20)	12(7-21)	165(25)	8(1.5-18)
On DMARDS	174(16)	-	322(29)	-	394(48)	-	366(56)	-	1256(34)
Psoriasis present	64(6)	-	74(7)	-	86(10)	-	58(9)	-	282(8)
Primary care referral	572(52)	22(2)	444(41)	18(2)	195(27)	9(1)	111(17)	3(0.4)	1322(36)
EIA	1022(93)	-	891(82)	-	632(77)	-	509(77)	-	3054(83)
1987 RA criteria *n (%)*	499(45)	-	406(37)	-	373(45)	-	294(45)	-	1572(43)
2010 RA criteria *n (%)*	656(60)	-	557(51)	-	453(55)	-	403(61)	-	2069(56)

IQR: interquartile range, HAQ: Health Assessment Questionnaire, DAS28: disease activity score with 28 joints, RF: rheumatoid factor, ACPA anti-citrullinated protein antibodies, CRP: C-reactive protein, DMARD: disease modifying anti-rheumatic drug, EIA: early inflammatory arthritis with symptom duration of less than 2 years *There were some missing data on date of symptom onset (see below), both age at onset and disease duration where derived from the date of symptom onset and date of birth and date of baseline assessment respectively. There were no missing data on either date of birth or date of baseline assessment.

3.1.2 Differences in baseline characteristics between patients who do and do not satisfy the 2010 ACR/EULAR classification criteria

Patients who satisfied the 2010 criteria at baseline were more likely to be female, were slightly older and had higher levels of disease activity, disability and CRP (table 3.1.2). Most markedly, ACPA or RF positivity was only present in a very small proportion of patients who did not fulfil the 2010 criteria (table 3.1.2). Interestingly disease duration at presentation was similar, if not slightly longer in the group who satisfied the criteria. In addition, the proportion of patients who had symptom duration of less than 2 years at baseline assessment was identical at 83%.

Table 3.1.2 Baseline characteristics of patients by 2010 criteria status

	2010 criteria positive n=2069	2010 criteria negative n=1603
Female *n(%)*	1422(69)	988(62)
Age onset *n(%)*	56(46-67)	53(40-66)
Symptom duration (weeks) *median(IQR)*	33(17-70)	30(16-68)
Smoking status *n(%)*		
current	538(26)	365(23)
previous	818(40)	596(37)
never	508(25)	402(25)
HAQ score *median(IQR)*	1.125(0.5-1.75)	0.5(0.125-1)
DAS28 *median(IQR)*	4.45(3.59-5.33)	2.72(2.12-3.42)
ACPA positive *n(%)*	639(31)	81(0.05)
RF positive *n(%)*	948(47)	162(0.1)
CRP *median(IQR)*	11(4-23)	6(0-13)
On DMARDs *n(%)*	1006(49)	549(34)
Psoriasis present *n(%)*	143(7)	139(9)
EIA *n(%)*	1723(83)	1331(83)
Satisfy 1987 criteria *n(%)*	1357(66)	216(13)

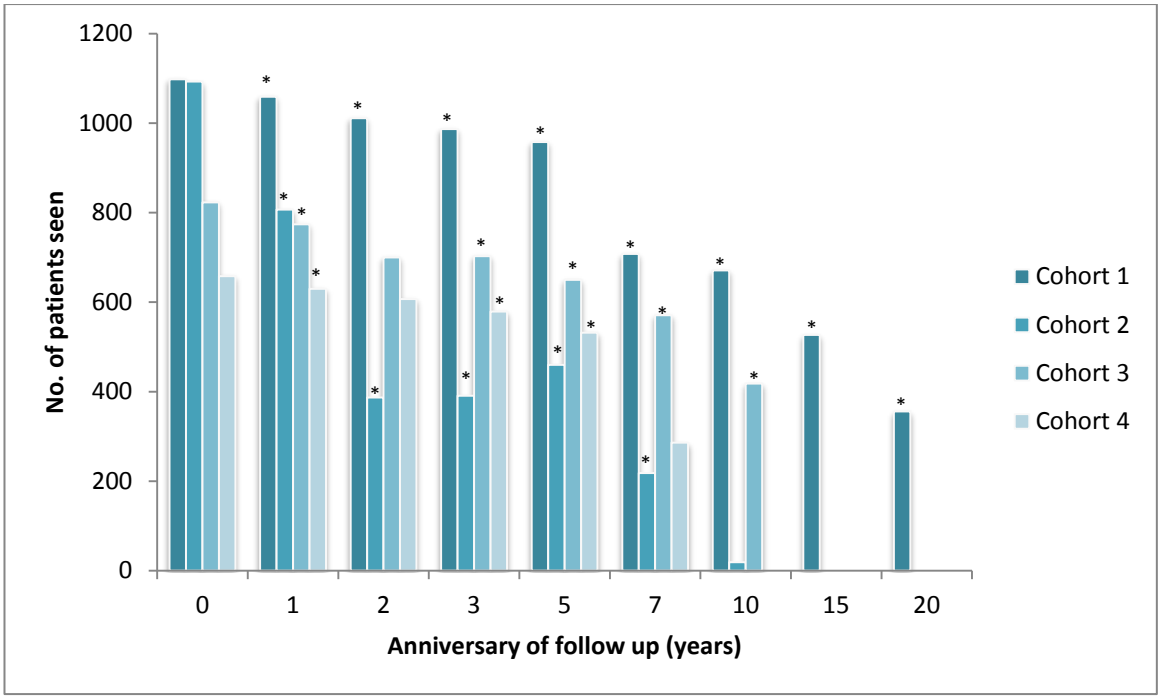
IQR: interquartile range, HAQ: Health Assessment Questionnaire, DAS28: disease activity score with 28 joints, RF: rheumatoid factor, ACPA anti-citrullinated protein antibodies, CRP: C-reactive protein, DMARD: disease modifying anti-rheumatic drug, EIA: early inflammatory arthritis with symptom duration of less than 2 years

3.2 Missing baseline data

Most of missing baseline data was due to patients not consenting to give blood samples (Table 3.1.1). For patients with missing data on ACPA within NOAR, search of contemporaneous medical records provided additional values for only 18 patients, as ACPA was rarely tested for as part of routine care prior to 2009. The other source of missing data was in the date of symptom onset. The majority of patients (3310, 90%) did not recall the exact day of the onset of their symptoms and for the purposes of data analysis this was imputed as the 15th day of the month A much smaller proportion (285, 0.08%) could not recall the month of their symptom onset and for these patients the month of June was used. There did not appear to be any significant differences in the proportions of missing data amongst the four cohorts, apart from ACPA status. The highest percentage of missing data on ACPA was within cohort 3. Patients who did not consent to give blood samples led to missing ACPA, RF and CRP data, however in addition there were changes in the laboratory practices which meant some of the samples obtained could not be tested for ACPA. As this is not related to the patients from whom the data were collected and testing was not done selectively by cohort, these missing data are likely to be MCAR.

3.3 Loss to follow up

Retention in the study by cohort is detailed in figure 3.3.1 below. In general loss to follow up was small, particularly in the first 5 years. After 5 years patients were only followed up further if they met certain pre-specified criteria detailed in section 2.1.7. Patients recruited in cohorts 3 and 4 appear to have significant drop off after 10 and 5 years respectively, this is due right censorship as the patients in those cohorts have not yet reached the point at which they would undergo the later follow up assessments.



*Follow should be complete in these cohorts to this time point.

Figure 3.3.1 Retention in NOAR cohorts over time

Deaths and embarkation

The ONS notified NOAR of 21 patients who had left the UK after recruitment, 13 from cohort 1, 4 from cohort 2, 3 from cohort 3 and 1 from cohort 4. Some of the apparent loss to follow up subsequently is due to patients dying. Deaths over time by cohort are shown below in figure 3.3.2. This has obviously impacted the earlier cohorts to a greater extent, and the loss to follow up not due to deaths is shown in figure 3.3.3. This shows the expected number of patients (based on the number recruited minus deaths by each time point) at each follow up alongside the number of patients actually seen. This has been restricted to the first 5 years as follow up should be complete in all cohorts and right censorship will not be relevant. Note that cohort 2 shows apparent significant loss to follow up, this was due to the limited scope of follow up planned for this cohort.

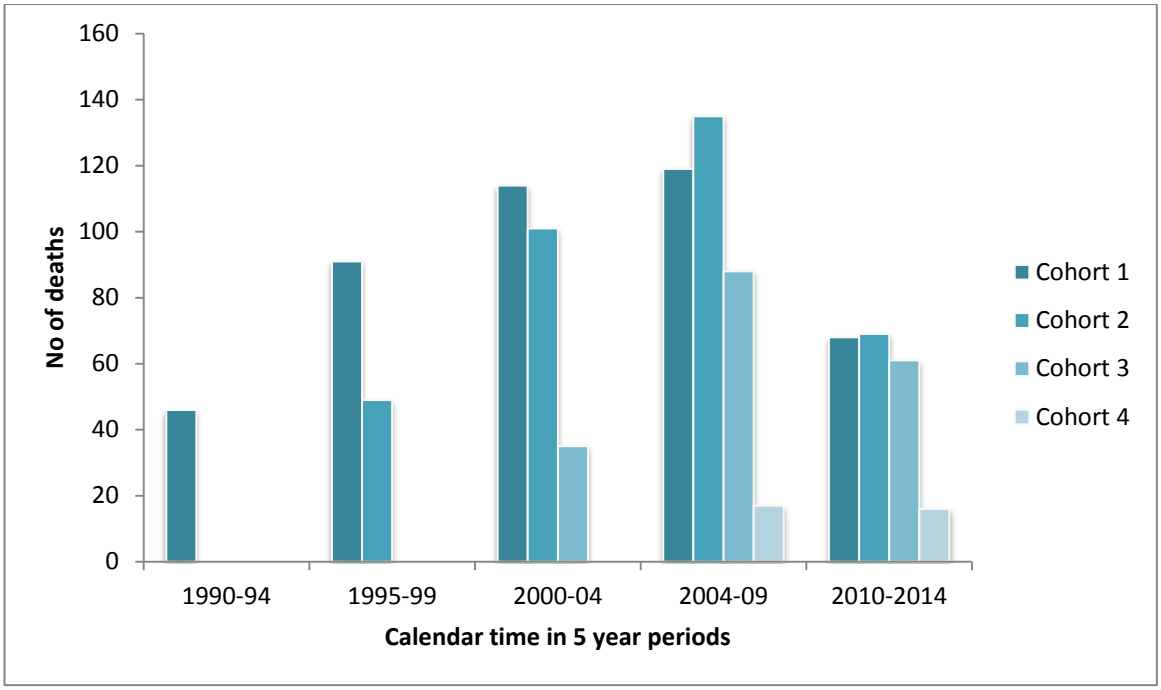
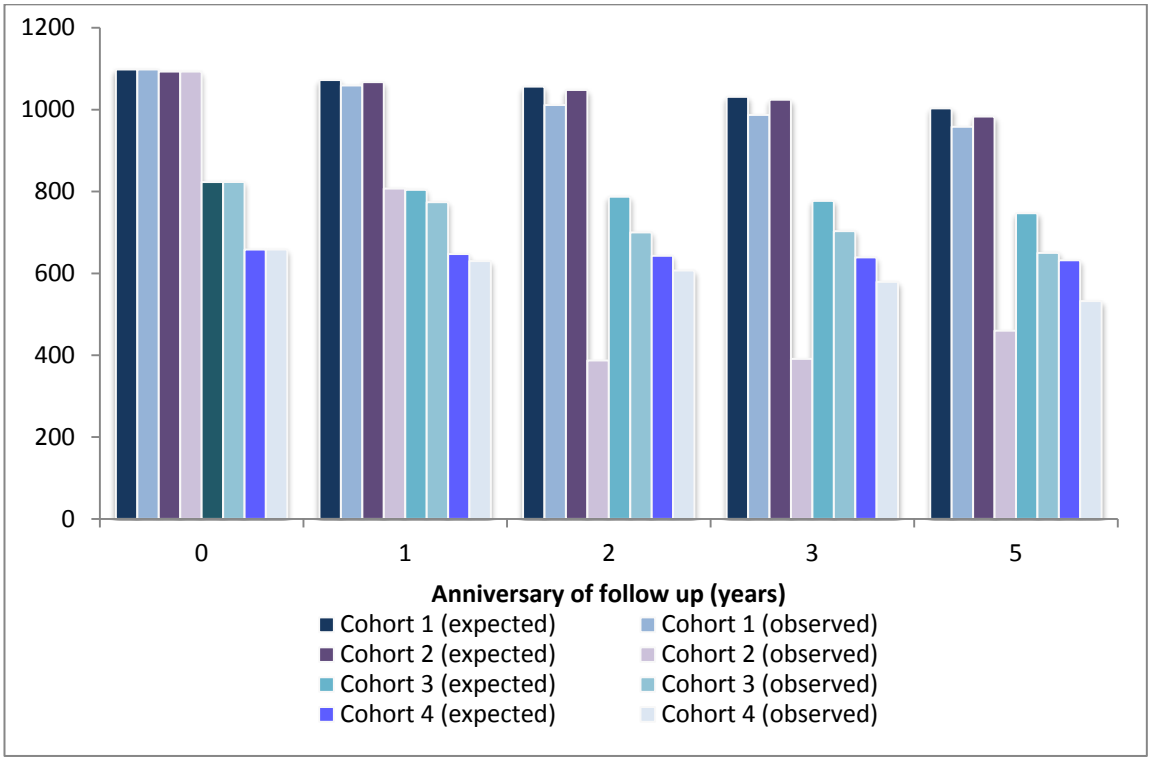


Figure 3.3.2 Deaths over time by cohort



Expected values were calculated by the subtracting the number of deaths occurring between follow ups from the number of patients expected at the previous follow up.

Figure 3.3.3 Expected and observed retention in NOAR cohorts over first 5 years of follow up.

3.4 Power

Retrospective power calculations were performed on the analyses in objectives 2 and 6 as examples.

The first example tested the hypothesis that patients with anti-CarP antibodies had more disability than patients without anti-CarP antibodies. With the 1995 patients included in that analysis, there was 90% power to detect a 7% difference in HAQ between the two groups (5% significance level).

The second example tested the hypothesis that patients who satisfy the 2010 criteria are more likely to die early. In the sample size of 1643 patients, there was again 90% power to detect a 4% difference in the number of deaths between patients who did and did not satisfy the criteria (5% significance level).

Chapter 4: Results

The incidence of RA

This chapter comprises one paper, in which the incidence of RA as defined by the 2010 ACR/EULAR classification criteria is calculated and compared to the incidence of RA as defined by the 1987 ACR criteria.

4 Results: The incidence of RA

4.1 The incidence of rheumatoid arthritis in the UK: comparisons using the 2010 ACR/EULAR classification criteria and the 1987 ACR classification criteria

Humphreys JH, Verstappen SM, Hyrich KL, Chipping JR, Marshall T, Symmons DP. *Annals of the Rheumatic Diseases* 2013 Aug; 72(8):1315-20

OPEN ACCESS

EXTENDED REPORT

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

Jennifer H Humphreys,[1] Suzanne M M Verstappen,[1] Kimme L Hyrich,[1] Jacqueline R Chipping,[2] Tarnya Marshall,[3] Deborah P M Symmons[1,4]

Handling editor Tore K Kvien

[1]Arthritis Research UK Epidemiology Unit, Manchester Academic Health Sciences Centre, University of Manchester, Manchester, UK
[2]Norfolk Arthritis Register, School of Medicine Health Policy and Practice, University of East Anglia, Norwich, UK
[3]Department of Rheumatology, Norfolk and Norwich University Hospital, Norwich, UK
[4]NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust and University of Manchester Partnership, Manchester, UK

Correspondence to
Professor Deborah PM Symmons, Arthritis Research UK Epidemiology Unit, Manchester Academic Health Sciences Centre, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, UK;
deborah.symmons@manchester.ac.uk

Accepted 7 August 2012
Published Online First 3 September 2012

ABSTRACT

Objectives The development of new classification criteria for rheumatoid arthritis (RA) calls for a re-estimation of RA incidence rates. The objectives of this study were to estimate the age and sex-specific incidence rates (IR) of RA in Norfolk, England using the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism criteria, and to compare those with IRs estimated using the 1987 ACR criteria.

Setting The Norfolk Arthritis Register (NOAR), a large primary care inception cohort of patients with inflammatory oligo- and polyarthritis (IP) aged ≥ 16.

Methods All patients notified to NOAR from 1990-5 with symptom onset in 1990 were included. The former Norwich Health Authority population was the denominator. Age and sex specific IRs using 1987 and 2010 classification criteria were calculated at baseline visit, annually for the first 3 years and at 5 years.

Results 260 patients were notified to NOAR with symptom onset in 1990 and without an alternative diagnosis. IRs applying the 2010 criteria at baseline were 54/100 000 for women and 25/100 000 for men. Age and sex-specific IRs using the 2010 classification criteria at baseline were similar to cumulative IRs applying the 1987 criteria up to 5 years. However, some patients only ever satisfied one set of criteria and a proportion of IA patients (20%) did not satisfy either criteria set over 5 years.

Conclusions The 2010 criteria classify similar numbers of patients as having RA at baseline, as the 1987 criteria would have taken up to 5 years to identify.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic, erosive inflammatory arthritis thought to affect approximately 1% of the UK adult population.[1] Recently it has been shown that aggressive early treatment can prevent much of the long term damage associated with the RA.[2] The 1987 American College of Rheumatology (ACR) classification criteria, widely used as entry criteria to clinical trials and observational studies, were developed in a cohort of patients with established, longstanding disease[3] and are known to perform poorly in patients presenting with recent onset inflammatory arthritis,[4] who may benefit most from early intensive treatment. The 2010 ACR/European League Against Rheumatism (EULAR) classification criteria for RA[5] aim to have improved sensitivity compared with the 1987 criteria. In particular, the 2010 criteria were designed to better identify RA in patients presenting soon after the development of signs and symptoms of the disease.

The developers of the new criteria describe them as 'defining a new paradigm of RA'. If this is the case, previous estimates of disease incidence and prevalence may no longer be accurate. Measuring prevalence in a relapsing remitting disease, or disease in which signs and symptoms resolve with treatment such as RA, presents additional challenges, as patients on treatment may be completely asymptomatic and have no signs of disease; therefore may be missed by population surveys. Measuring incidence requires an inception cohort with complete capture of all new cases of disease within a stable, defined background population. To date, no studies have estimated incidence of RA using the 2010 criteria. The objectives of this study were (i) to estimate age and sex-specific incidence of RA using the 2010 criteria in Norfolk, UK and (ii) to compare these incidence rates (IR) with those using the previous criteria set, at initial presentation and cumulatively over 5 years.

PATIENTS AND METHODS

Setting

The Norfolk Arthritis register (NOAR) is a primary care inception cohort of patients aged ≥16 years presenting with ≥2 swollen joints for at least 4 weeks to either primary or secondary care within the former Norwich Health Authority. A detailed description of NOAR is available in previous publications.[6] Briefly, patients undergo standardised assessment by a research nurse including details of symptom onset, 51 swollen and tender joint counts and examination for nodules, as well as consent to medical records review. Assessments (including joint counts and examination for nodules) are repeated annually for the first 3 years and at 5 years. Blood is taken at baseline and after 5 years for C reactive protein (CRP) and rheumatoid factor (RF) (latex test) and the remaining sera stored frozen; this was subsequently used to measure anti-citrullinated protein antibody status (ACPA) (Axis-Shield Diastat Anti-CCP kit, Dundee,

Scotland). Patients included in this current analysis were all those who had symptom onset of joint pain or swelling between January and December 1990, and were notified to NOAR within 5 years of symptom onset. This time period was selected as we can only be reasonably sure that all new cases of inflammatory oligo- and polyarthritis (IP) presenting to primary care were identified in 1990–1994 and we have previously reported on the incidence using the 1987 criteria for patients with a symptom onset in 1990.[6 7] It should be noted that a group of patients from NOAR were used in the development of the 2010 criteria.[5] However, those patients were recruited since 2000 and none of the patients included in the present study formed part of the criteria development cohort.

Application of classification criteria

For the 2010 criteria, joint counts and duration of symptoms were obtained from the nurse assessments and weighted scores assigned as detailed in the criteria (figure 1). In order to obtain as complete a dataset as possible, the medical records of those patients included in this analysis who did not provide a blood sample were searched to identify acute phase reactant (CRP or erythrocyte sedimentation rate (ESR)) and RF results taken near to the time of symptom onset. CRP and ESR were considered elevated if >5 mg/l and >10 mm/h respectively, according to local laboratory reference ranges. The 2010 criteria divide values of RF and ACPA into the following groups for scoring: negative: defined as ≤ upper limit of normal (ULN) for the laboratory and assay; low positive: >ULN but ≤3 times ULN; and high positive: >3 times ULN. In this study these cut offs were 40 and 120 International Units (IU) respectively for RF; for ACPA they were 5 and 15 IU respectively.

The 1987 criteria exist in two formats: list (figure 1) and tree.[3] The list format includes radiographic erosions, which can be substituted with clinical data in the tree format. At baseline assessment, radiographs were not taken, thus the tree format was applied; at 5 years all patients underwent radiographic examination of hands and feet and patients were said to have met the 1987 criteria if they satisfied the tree or list format.

For both criteria sets, if data were missing on any variables, total scores were calculated with the missing variable value taken as zero, and patients said to have met the criteria if they reached the defined cut-offs: ≥6/10 for the 2010 criteria and ≥4/7 for the 1987 criteria.

Incidence rates

The denominator population was provided by the former Norfolk Health Service Authority.[6] Both criteria sets were applied to calculate age and sex-specific IRs at the baseline assessment. Using the 2010 criteria, 5 year cumulative incidence was estimated by taking the highest score for each parameter (joint count, serology, acute phase reactants and symptom duration) at any assessment over the 5 years follow up period. For the 1987 criteria, 5 years cumulative incidence was estimated in the following manner: if a patient satisfied a particular criterion at an individual assessment, it was then carried forward to all future assessments. CIs around the IRs were calculated using the Poisson distribution.

NOAR is approved by the Norwich Local Research Ethics Committee and all patients gave written consent. All data were analysed using STATA V.10 software package (Stata, College Station, Texas, USA).

RESULTS

A total of 283 patients were registered with NOAR who had symptom onset in 1990. Of these, 23 patients were diagnosed with other rheumatological disorders by their treating rheumatologist and were therefore excluded. Table 1 shows baseline demographic data of the cohort and the proportion of missing data. Thirty-six patients declined to provide a blood sample at baseline. Despite medical record review, 31 of these patients had no result for acute phase reactants and 12 patients had no autoantibody results at baseline. five patients continued to decline blood sampling throughout follow up and therefore had no results available for the acute phase reactant or either autoantibody parameters. After 5 years, 25 patients had died, 22 patients declined follow up after baseline assessment and 16 patients were lost to follow up, thus a total of 197 patients remained under active follow up. For the cumulative analysis, patients who did not complete 5 years follow up were classified cumulatively up to their last assessment.

The overall IR when applying the 2010 criteria at baseline was 40/100 000; 54/100 000 for women and 25/100 000 for men. These rates were higher than when applying the 1987 criteria at baseline (32/100 000 overall, 45/100 000 for women and 18/100 000 for men). Age and sex-specific IRs using the 2010 classification criteria at baseline showed marked similarities to cumulative IRs applying the 1987 criteria up to 5 years (table 2). In women the peak age of incidence was younger than in men for both criteria sets, with highest rates between ages 45–74. In men incidence appeared to increase with age with highest rates in men over 65 years old.

Applying the 2010 criteria cumulatively over 5 years follow up gave an estimated IR of 48/100 000; for the 1987 criteria this was 44/100 000. A further 34 patients satisfied the 2010 criteria when applied cumulatively over 5 years; applying the 1987 criteria cumulatively for 5 years classified 49 additional patients as RA. Results applying both criteria sets cumulatively converged after approximately 3 years follow up (figure 2 and table 3); nevertheless there remained some discordance between the criteria (table 4). After 5 years follow up, 50 (19%) patients satisfied neither criteria set, cumulatively or cross-sectionally. All five patients who had no blood results throughout the follow up period met at least one criteria set at baseline.

DISCUSSION

The 2010 ACR/EULAR classification criteria for RA have provided a new definition for the disease entity 'RA'. This is the first study to estimate the incidence of RA using the 2010 criteria. We have shown, in a cohort of patients with early IP, that the incidence of RA according to the 2010 criteria is higher at baseline assessment than the incidence of RA according to the 1987 criteria. The 2010 criteria appear to identify at baseline similar rates of RA as the 1987 criteria identify cumulatively over 5 years. We have shown previously that cumulative application of the 1987 criteria over 5 years increases incidence estimates by up to 93%.[7] However, this requires long term follow up of all patients presenting with undifferentiated inflammatory arthritis. Today, with improved treatment strategies, some patients who are given a clinical diagnosis of RA by their treating physicians may never satisfy them. Our results show that application of the 2010 criteria in early disease may therefore negate the need for such long term follow up to confirm classification, and may in part address concerns that some patients, whose disease is suppressed by appropriate treatment, may be inappropriately classified as not having RA.

Criterion	Definition
A patient is classified as RA if 4/7 criteria are satisfied. Criteria 1-4 must have been present for ≥6 weeks	
1. Morning stiffness	Morning stiffness in and around the joints, lasting at least an hour before maximal improvement
2. Arthritis of ≥3 joints areas	≥3 joints areas simultaneously have had synovitis observed ay a physician
3. Arthritis of hand joints	At least 1 area swollen in a wrist, MCP or PIP joint
4. Symmetric arthritis	Simultaneous involvement of the same joint areas on both sides of the body
5. Rheumatoid nodules	Subcutaneous nodules, over bony prominences, extensor surfaces or juxta-articular regions
6. Serum rheumatoid factor (RF)	Positive RF
7. Radiographic changes	Radiographic changes typical of RA in posteroanterior hand and wrist radiographs

1987 ACR Classification criteria for RA[3]

Target population: Patients who (i) have at least one joint with clinical synovitis, and (ii) the synovitis not better explained by another disease	Score
Add score of categories A-D, score of ≥6/10 needed to classify patient as having definite RA	
A. Joint involvement (tender/swollen)	
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	
Negative RF /ACPA	0
Low-positive RF/low positive ACPA	2
High positive RF/high-positive ACPA	3
C. Acute phase reactants	
Normal CRP&ESR	0
Abnormal CRP/ESR	1
D. Duration of symptoms	
<6 weeks	0
≥6 weeks	1

2010 ACR/EULAR Classification criteria for RA[5]

Figure 1 Classification criteria for RA.

Incidence of RA has been estimated in a variety of populations, with considerable variation in the results.[8] In the USA, the incidence of RA in Olmsted County, Minnesota has been tracked since 1955 using the Rochester Epidemiology Project medical record linkage system.[9] They, and others, have shown a decline of the incidence of RA in the second half of the 20th century.[10 11] Interestingly, this trend may have slowed or even reversed in the past 10 years and their latest published IR was 41/100 000 population,[12] which is higher than our estimate using the 1987 criteria. It will be interesting to assess whether these long term trends in incidence will continue given the re-definition of disease in the 2010 criteria. Another recent estimation of RA incidence based on the 1987 criteria was undertaken in Spain, where cases were identified from primary care during the establishment of a nationwide programme of early arthritis units.[13] They estimated an IR of 8/100 000, significantly lower than ours applying the same criteria set. This may be due to the reported lower incidence of RA in Southern Europe compared to Northern Europe (8). Where inception cohorts are not available, other methods have been used to estimate incidence. In the UK, a combination of diagnostic codes and disease modifying drug prescriptions recorded within the General Practice Research Database (GPRD)[14] were used to identify new cases; in Finland insurance claim forms have been used.[11] In both cases, data were collected retrospectively, and, in particular with the GPRD, the definition of incident RA is vulnerable to misclassification.

Table 1 Baseline demographics and criteria variables

Demographic	Frequency	Missing n (%)
Age at symptom onset (mean (SD))	54 (16.2)	0
Female (n (%))	173 (69)	0
Symptom duration in weeks (median (IQR))	29.6 (4.3–71.9)	0
RF low positive (n (%))	25 (11)	28 (11)†
High positive (n (%))	47 (20)	
ACPA low positive (n (%))	6 (3)	77 (30)†
High positive (n (%))	38 (20)	
Joint involvement* (n (%))		0
1 large joint	9 (3)	
2–10 large joints	9 (3)	
1–3 small joints	41 (16)	
4–10 small joints	52 (20)	
>10 joints	149 (57)	
Acute phase reactant positive (n (%))	120 (52)	27 (10)
CRP	116 (48)	
ESR	9 (64)	
CRP (mean (std dev))	19 (35)	
ESR (mean (std dev))	30 (34)	
Morning stiffness ≥60 min (n (%))	172 (66)	0
Arthritis of ≥3 joints areas (n (%))	172 (66)	0
Arthritis of hand joints (n (%))	215 (83)	0
Symmetric arthritis (n (%))	183 (70)	0
Rheumatoid nodules (n (%))	19 (7)	0

*Large joints were defined as shoulders, elbows, hips, knees and ankles. Small joints are defined as metocarpophalangeal joints, proximal interphalageal joints, second through fifth metatarsophalangeal joints, thumb interphalangeal joints and wrists. Distal interphalangeal joints, first carpometacarpal joints and first metatarsophalangeal joints were excluded as per the 2012 criteria.[5]
†Missing data quoted are for individual autoantibodies. 8 (2%) patients had no results for ACPA or RF.
ACPA, anti-citrullinated protein antibody, CRP, C reactive protein; ESR, erythrocyte sedimentation rate; RF, rheumatoid factor.

Studies assessing the 2010 criteria to date have mainly focused on their sensitivity and specificity to predict surrogates of an RA diagnosis (for which there is no gold standard) such as initiation of disease modifying anti-rheumatic drug therapy,[15 16] physician opinion[17] and absence of drug free remission.[18] These studies have shown that the 2010 criteria classify more patients as RA earlier in the disease course compared to the 1987 criteria, with a general improvement in sensitivity at the cost of specificity. Our findings support this hypothesis, that the 2010 criteria are better at classifying early RA; in addition we have demonstrated that this earlier classification identifies similar rates of disease. However, the lack of gold standard is also a limitation in our study, as without this it is not possible to measure true incidence.

This study highlights certain subgroups of patients who may be of particular interest for further investigation. The first is those patients who only met one criteria set over the 5 years

Table 2 Age and sex specific incidence rates (IR/100 000 population)

		Female patients		Male patients	
Age band	No. of patients with inflammatory oligo- and polyarthritis	2010 criteria at baseline IR (95% CI)	1987 criteria cumulative to 5 years follow up IR (95% CI)	2010 criteria at baseline IR (95% CI)	1987 crtieria cumulative to 5 years follow up IR (95% CI)
15–24	17	18.6 (6.8 to 40.6)	15.5 (5.0 to 36.3)	0 (0 to 11.1)	6 (0.7 to 21.7)
25–34	23	20.3 (8.2 to 41.8)	31.9 (15.9 to 57.0)	5.6 (0.7 to 20.3)	8.4 (1.7 to 24.6)
35–44	34	56.6 (34.1 to 88.4)	56.6 (34.1 to 88.4)	12.1 (3.3 to 30.9)	12.1 (3.3 to 30.9)
45–54	53	85.6 (56.4 to 124.6)	98.3 (66.8 to 139.5)	34.5 (17.2 to 61.7)	31.4 (15.0 to 57.7)
55–64	58	91.8 (59.4 to 135.5)	91.8 (59.4 to 135.5)	42.1 (21.0 to 75.3)	42.1 (21.0 to 75.3)
65–74	53	87.1 (55.8 to 129.6)	94.4 (61.7 to 138.3)	58.3 (31.9 to 97.8)	66.6 (38.1 to 108.2)
75+	22	26.1 (10.5 to 53.7)	29.8 (12.9 to 58.7)	44.3 (17.8 to 91.3)	57.0 (26.1 to 108.1)
Total	260	53.9 (44.5 to 64.7)	58.5 (48.7 to 69.8)	24.5 (18.1 to 32.4)	27.5 (20.7 to 35.8)

follow up period. The characteristics of these patients reflect the criteria themselves: patients satisfying only the 1987 criteria were more likely to have prolonged morning joint stiffness, they also had more symmetrical and more hand joint involvement. By contrast, patients satisfying only the 2010 criteria had greater number of joints involved at each assessment (reflecting the inclusion of tender as well as swollen joints). The most notable difference is seen in the frequency of autoantibodies. The majority of patients in our cohort who never satisfied the 2010 criteria but did satisfy the 1987 criteria were autoantibody negative; this difference was most marked at baseline assessment. This pattern has been noted in other cohorts,[19 20] and it has been postulated that the two criteria sets may be describing different clinical entities.[21] However, the striking similarity in IRs over time in our patients argues against this. It may be that the two criteria sets represent different aspects of the same disease construct; the 2010 criteria describe an acute inflammatory arthritis, whereas the 1987 criteria describe the long term damage that occurs as a consequence. Another subgroup of interest is those patients who never satisfy either criteria set. For patients remaining in this study 5 years after symptom onset, this group comprised 50 patients (19%), which is a substantial proportion. There were missing data in our cohort, particularly relating to serological markers, and this may have led to some patients being misclassified as non-RA. However, none of the patients who could not be classified by either criteria set over 5 years had missing data on all serological variables at all time points. Investigating the long term outcomes of the patients who satisfy neither criteria set will be important to assess the validity of the 2010 criteria.

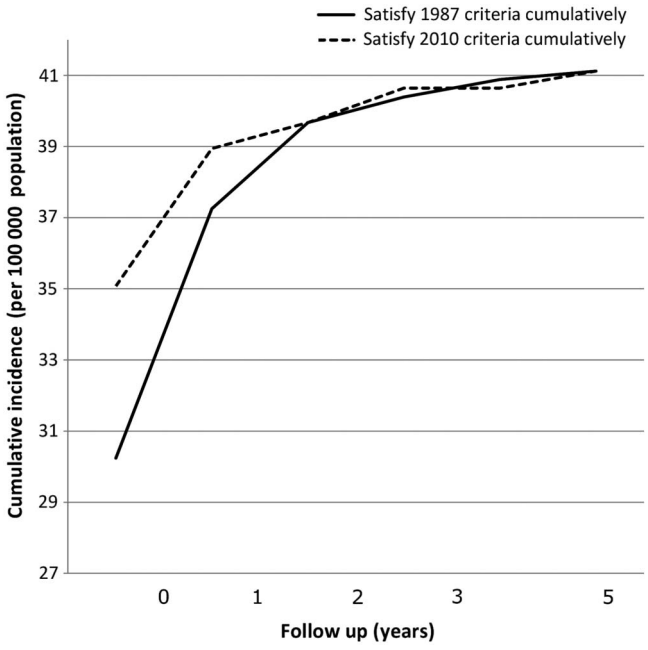


Figure 2 Cumulative incidence of RA in patients satisfying both criteria sets after 5 years (n=170).

In the publication describing the development of the 2010 criteria,[5] and in subsequent editorials[22] the authors suggest they may be used in clinical practice to allow access to disease modifying anti-rheumatic drug or biologic therapy. Although we have shown that these criteria classify more patients early in the disease course than the previous criteria set (which were never used in this context), our results suggest they are not sufficiently sensitive for this purpose. In particular, the fact that some patients fulfil the previous criteria set without fulfilling the new criteria, even after 5 years follow up, indicate caution should be taken considering this application. Further work is needed to elucidate the long term outcomes of patients not fulfilling the new criteria to answer this question. If these are universally good for all patients not fulfilling the criteria, their use as a gateway to treatment may be appropriate. There are a number of strengths in the present study due to unique features of NOAR in the UK: Norfolk has a stable population with little migration, there is a balanced mix of rural and urban populations (thus is representative of both) and a central referral system for musculoskeletal patients to a single secondary care provider, Norwich and Norfolk University Hospital. Significant efforts were made to ensure all patients with IP newly presenting to primary care were reported to NOAR

Table 3 Patients satisfying rheumatoid arthritis criteria cumulatively over time

	Satisfy 1987 criteria cumulatively	Satisfy 2010 criteria cumulatively	Satisfy both criteria sets cumulatively	Satisfy 1987 criteria cumulatively if satisfy both by 5 years	Satisfy 2010 criteria cumulatively if satisfy both by 5 years
Baseline	131/260	166/260	119/260	125/170	145/170
1 year	163/260	186/260	150/260	154/170	161/170
2 years	174/260	193/260	159/260	164/170	164/170
3 years	177/260	197/260	165/260	167/170	168/170
5 years	180/260	200/260	170/260	170/170	170/170

Table 4 Number of patients satisfying each criteria set after 5 years follow up

	Patients satisfying 1987 criteria cumulatively *n (%)*	Patients not satisfying 1987 criteria cumulatively *n (%)*	Total
Patients satisfying 2010 criteria cumulatively	170 (65)	30 (12)	200
Patients not satisfying 2010 criteria cumulatively	10 (4)	50 (19)	60
Total	180	80	260

when it was first established in 1989, with visits to GP practices, advertising and a small incentive. We therefore selected the year 1990 to estimate incidence in this study as the year with near complete capture of all patients presenting with early IP. Nevertheless, the IRs reported here are likely to be an underestimate for a number of reasons. Some patients only had RF or ACPA measured; a high positive result in the other autoantibody may have increased the number of patients classified as RA. However, the 2010 criteria only require testing of either RF or ACPA, therefore our data represent a valid estimate of incidence. NOAR was established when the 1987 criteria were the standard for classification, and elements of its design may make classification by the 1987 criteria easier than by the 2010 criteria, potentially reducing the IRs using the 2010 criteria. This highlights difficulties that occur when applying criteria retrospectively to an historic cohort. In addition, there may have been cases which were not captured, including patients who did not seek healthcare advice at the time of symptom onset. To allow for this delay and to obtain as true an estimate of incidence in that year as possible, the age and sex-specific IRs reported at baseline included patients who had presented to NOAR up to 5 years after symptom onset. However, this meant that a small number of patients had been symptomatic of their disease for a number of years at the time of initial assessment. If IRs were calculated based on initial assessments of only those patients who presented within 2 years of symptom onset, the overall IR using the 2010 criteria was 35/100 000 population; for 1987 criteria it was 27/100 000 at baseline presentation but increased to 36/100 000 cumulatively 5 years after symptom onset.

A further limitation relates to erosive disease. The 2010 criteria include an amendment which states that any patient with radiological evidence of erosion typical of RA should automatically be classified as having RA, without the need to fulfil any other aspect of the criteria. As radiographs were not performed at baseline in this cohort, and because there is no clear definition of 'typical RA erosion', this was not applied in the present analysis. x-Rays were performed on all patients after 5 years follow up; if the presence of any erosion (although not specifically a 'typical RA' erosion) was applied at that point, four further patients (three women and one man) could be classified as having RA according to the 2010 criteria.

In conclusion, we have reported the first IR estimates of RA applying the 2010 ACR/EULAR classification criteria. We have shown that the incidence of RA, as estimated by the 2010 classification criteria at baseline, is very similar to the estimates using the 1987 criteria cumulatively over 5 years. These results indicate that the 2010 criteria may identify RA patients earlier in the disease course and will be important in order to plan timely, cost-effective and efficacious management of patients presenting with inflammatory arthritis.

Acknowledgements The authors gratefully acknowledge the support of clinical staff at the Norfolk and Norwich University Hospital, local primary care physicians and all of the Norfolk Arthritis Register research nurses. Data management by the team in Manchester is also appreciated.

Contributors All authors were involved in drafting the article or critically reviewing and revising it, and all authors approved the final version to be published. Study concept and design: JHH, KLH, DPMS; acquisition of data: JRC, TM; analysis and interpretation of the data: JHH, SMMV, KLH, DPMS.

Funding NOAR is funded by Arthritis Research UK (grant reference 17552). JHH is funded by an Arthritis Research UK Clinical fellowship (grant reference 19743). The authors have received no other financial support or other benefits from commercial sources for the work reported in the manuscript.

Competing interests None.

Ethics approval Norwich Local Research Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Symmons D,** Turner G, Webb R, *et al*. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology (Oxford)* 2002;**41**:793–800.
2. **Goekoop-Ruiterman YP,** de Vries-Bouwstra JK, Allaart CF, *et al*. Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2007;**146**:406–15.
3. **Arnett FC,** Edworthy SM, Bloch DA, *et al*. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;**31**:315–24.
4. **Harrison BJ,** Symmons DP, Barrett EM, *et al*. The performance of the 1987 ARA classification criteria for rheumatoid arthritis in a population based cohort of patients with early inflammatory polyarthritis. American Rheumatism Association. *J Rheumatol* 1998;**25**:2324–30.
5. **Aletaha D,** Neogi T, Silman AJ, *et al*. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European league against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;**69**:1580–8.
6. **Symmons DP,** Barrett EM, Bankhead CR, *et al*. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk arthritis register. *Br J Rheumatol* 1994;**33**:735–9.
7. **Wiles N,** Symmons DP, Harrison B, *et al*. Estimating the incidence of rheumatoid arthritis: trying to hit a moving target? *Arthritis Rheum* 1999;**42**:1339–46.
8. **Alamanos Y,** Voulgari PV, Drosos AA. Incidence and prevalence of rheumatoid arthritis, based on the 1987 American College of Rheumatology criteria: a systematic review. *Semin Arthritis Rheum* 2006;**36**:182–8.
9. **Maradit KH,** Crowson CS, Gabriel SE. Rochester Epidemiology project: a unique resource for research in the rheumatic diseases. *Rheum Dis Clin North Am* 2004;**30**:819–34, vii.
10. **Doran MF,** Pond GR, Crowson CS, *et al*. Trends in incidence and mortality in rheumatoid arthritis in Rochester, Minnesota, over a forty-year period. *Arthritis Rheum* 2002;**46**:625–31.
11. **Kaipiainen-Seppanen O,** Kautiainen H. Declining trend in the incidence of rheumatoid factor-positive rheumatoid arthritis in Finland 1980–2000. *J Rheumatol* 2006;**33**:2132–8.
12. **Myasoedova E,** Crowson CS, Kremers HM, *et al*. Is the incidence of rheumatoid arthritis rising?: results from Olmsted County, Minnesota, 1955–2007. *Arthritis Rheum* 2010;**62**:1576–82.
13. **Carbonell J,** Cobo T, Balsa A, *et al*. The incidence of rheumatoid arthritis in Spain: results from a nationwide primary care registry. *Rheumatology (Oxford)* 2008;**47**:1088–92.
14. **Rodriguez LA,** Tolosa LB, Ruigomez A, *et al*. Rheumatoid arthritis in UK primary care: incidence and prior morbidity. *Scand J Rheumatol* 2009;**38**:173–7.
15. **Alves C,** Luime JJ, Van ZD, *et al*. Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). *Ann Rheum Dis* 2011;**70**:1645–7.

16. **Cader MZ,** Filer A, Hazlehurst J, *et al*. Performance of the 2010 ACR/EULAR criteria for rheumatoid arthritis: comparison with 1987 ACR criteria in a very early synovitis cohort. *Ann Rheum Dis* 2011;**70**:949–55.
17. **Britsemmer K,** Ursum J, Gerritsen M, *et al*. Validation of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: slight improvement over the 1987 ACR criteria. *Ann Rheum Dis* 2011;**70**:1468–70.
18. **van der Linden MP,** Knevel R, Huizinga TW, *et al*. Classification of rheumatoid arthritis: comparison of the 1987 American College of Rheumatology criteria and the 2010 American College of Rheumatology/European league against Rheumatism criteria. *Arthritis Rheum* 2011;**63**:37–42.
19. **Kaneko Y,** Kuwana M, Kameda H, *et al*. Sensitivity and specificity of 2010 rheumatoid arthritis classification criteria. *Rheumatology (Oxford)* 2011;**50**:1268–74.
20. **Fautrel B,** Combe B, Rincheval N, *et al*. Level of agreement of the 1987 ACR and 2010 ACR/EULAR rheumatoid arthritis classification criteria: an analysis based on ESPOIR cohort data. *Ann Rheum Dis* 2012;**71**:386–9.
21. **De Hair MJH,** Lehmann KA, Van De Sande MGH. The 2010 ACR/EULAR classification criteria for rheumatoid arthritis: Earlier diagnosis at the expense of increased heterogeneity. Arthritis Rheum Conference, October.
22. **Combe B.** The new classification criteria for rheumatoid arthritis and their impact on therapeutic decisions. *Joint Bone Spine* 2011;**78**:539–41.

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

Jennifer H Humphreys, Suzanne M M Verstappen, Kimme L Hyrich, Jacqueline R Chipping, Tarnya Marshall and Deborah P M Symmons

Ann Rheum Dis 2013 72: 1315-1320 originally published online September 3, 2012
doi: 10.1136/annrheumdis-2012-201960

4.1.1 Supplementary results

Erratum: Table 1 in paper 4.1 shows 27 patients with missing acute phase reactants at baseline. In line with the result reported in the text, this should read 31.

4.1.2 Differences between patients with missing and non-missing baseline data

Table 4.1.1 shows the characteristics of patients with and without missing blood samples in the analysis in chapter 4. Table 4.1.2 shows the results of the univariate linear or logistic regression analysis testing for differences between patients with missing data and those with complete data. The proportion of patients who fulfilled the 1987 criteria at baseline was smaller in the missing group, and they had longer disease duration at presentation. Neither of these were significant associations in the regression models.

Table 4.1.1 Characteristics of patients with missing blood sample at baseline compared to non-missing

	Missing *n=31*	Non-missing *n=229*
Female	22(71)	156(68)
Age onset	54(40-62)	56(43-67)
Symptom duration *(weeks)*	43(14-112)	30(14-66)
Satisfy 1987 criteria at baseline	11(35)	120(52)
Satisfy 2010 criteria at baseline	20(65)	147(64)

Table 4.1.2 Univariate linear/logistic regression testing for differences between patients with missing baseline bloods compared to non-missing

	OR/β* (95% CI)
Female	1.14(0.50, 2.61)
Age onset	-3.70(-9.99, 2.59)
Symptom duration *(weeks)*	160(-0.86, 322)
Satisfy 1987 criteria at baseline	0.50(0.23, .1.09)
Satisfy 2010 criteria at baseline	1.01(0.46,2.22)

*Beta coefficient for continuous variables, odds ratio (OR) for binary variables

Chapter 5: Results

Mortality and RA

This chapter comprises three papers, all of which investigate different aspects of the increased mortality associated with RA. The first investigates the ability of the 2010 classification criteria to predict mortality in patients with early inflammatory arthritis (EIA). The second paper examines whether the levels or number of the two autoantibodies which form one item of the 2010 criteria (RF and ACPA) affect mortality risk in patients with EIA. In the final paper, the 2010 criteria are used to define RA in a study comparing mortality rates in patients with inflammatory arthritis to rates in the general population over 20 years.

5 Results: Mortality and RA

5.1 2010 ACR/EULAR classification criteria for rheumatoid arthritis predict increased mortality in patients with early arthritis: results from the Norfolk Arthritis Register (NOAR)

Humphreys JH, Verstappen SM, Hyrich KL, Chipping JR, Symmons DP. *Rheumatology* 2013; 52(6):1141-2

RHEUMATOLOGY

Letters to the Editor

Rheumatology 2013;52:1141–1142
doi:10.1093/rheumatology/ket113
Advance Access publication 5 March 2013

2010 ACR/EULAR classification criteria for rheumatoid arthritis predict increased mortality in patients with early arthritis: results from the Norfolk Arthritis Register

SIR, The absence of a gold standard for diagnosis in RA has led to the development of classification criteria, primarily for the purpose of standardization of entry into clinical trials and other clinical studies. Previously, increased mortality was identified in patients defined as having RA by the 1987 ACR classification criteria [1]. In recent years, a consensus within the rheumatology community has emerged for early initiation of treatment in RA. The 1987 criteria are insufficiently sensitive at this point in the natural history of the disease. To address this, new classification criteria were published in 2010 [2] and have been shown to identify a greater proportion of patients as having RA when they first present [3–6]. It is important that the 2010 criteria are validated with respect to all previously recognized adverse outcomes associated with RA, including mortality. Therefore, we aimed to examine whether, in a cohort of patients with early inflammatory polyarthritis, the 2010 ACR/EULAR classification criteria for RA identified those with decreased survival, and how they compared with the 1987 criteria.

The study was set in the Norfolk Arthritis Register (NOAR), a primary care inception cohort of adults (aged $\geqslant$16 years) with inflammation in $\geqslant$2 joints for $\geqslant$4 weeks, recruited between 1990 and 2009 [7]. Patients were included in this analysis if they had symptom duration of $<$2 years at baseline and had not received DMARD therapy prior to initial assessment. At the first visit, patients were assessed by a research nurse who performed a joint examination (51 tender and swollen joint counts), applied the tree format of the 1987 criteria and took blood; sera were stored frozen and analysed for CRP, RF and subsequently, ACPAs. The 2010 criteria were applied retrospectively using data from the baseline visit. All patients registered to NOAR are flagged with the Office for National Statistics (ONS); NOAR is notified of any deaths with a lag in reporting of $\sim$6 months. All deaths prior to 30 June 2012 were included to allow for this lag. The risk of early death was modelled using Cox proportional hazard estimation univariately for each criteria set, then adjusting for age and sex. All data were analysed using STATA 11 software package (Stata, College Station, TX, USA). NOAR is approved by the Norfolk Local Ethics Committee and all patients gave written consent.

Complete data were available for 1643 patients, with 20 700 person-years follow-up. The median (interquartile range) age at symptom onset was 55 (43–68) years and 1074 (65%) patients were female. At baseline, significantly more patients satisfied the 2010 criteria, 892 (54%), than the 1987 criteria, 658 (40%) (Pearson's $\chi^2 = 764$, $P < 0.001$). The ONS reported 466 deaths (28%) over the follow-up period. In the unadjusted Cox proportional hazard model, patients who fulfilled the 2010 criteria had a significantly increased risk of early death compared with those patients in NOAR who did not fulfil these criteria, and the association was maintained in the age- and sex-adjusted model, HR 1.35 (95% CI 1.13, 1.64). Similar results were seen with the 1987 criteria, although the age- and sex-adjusted model identified a lower level of increased risk than the 2010 criteria, HR 1.24 (1.03, 1.49); this may be because patients who fulfilled the 1987 criteria were, on average, older than patients who fulfilled the 2010 criteria (mean age 55 *vs* 53 years).

This is the first study to show that, in patients presenting with early inflammatory arthritis, those who fulfil the 2010 classification criteria for RA have significantly increased mortality compared with those who do not. The 2010 criteria appear to be as efficient as the 1987 criteria in identifying this increased risk of mortality. Further, they identify a greater proportion of at-risk patients soon after their first presentation to health care. This study further validates the 2010 criteria in their ability to identify early those patients with inflammatory arthritis at risk of poor outcomes.

Rheumatology key message

- Mortality is increased in patients with early arthritis who fulfil the 2010 criteria for RA.

Funding: NOAR is funded by Arthritis Research UK (grant reference 17552). J.H.H. is funded by Arthritis Research UK (grant reference 19743).

Disclosure statement: The authors have declared no conflicts of interest.

Jennifer H. Humphreys[1], Suzanne M. M. Verstappen[1], Kimme L. Hyrich[1], Jacqueline Chipping[2] and Deborah P. M. Symmons[1,3]

[1]*Arthritis Research UK Epidemiology Unit, Manchester Academic Health Sciences Centre, University of Manchester, Oxford Road, Manchester, UK,* [2]*Norfolk Arthritis Register, School of Medicine Health Policy and Practice, University of East Anglia, Norwich, UK and* [3]*NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust and University of Manchester Partnership, Manchester, UK.*
Accepted 4 February 2013
Correspondence to: Deborah P. M. Symmons, Arthritis Research UK Epidemiology Unit, Manchester Academic Health Sciences Centre, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, UK.
E-mail: deborah.symmons@manchester.ac.uk

References

1 Arnett FC, Edworthy SM, Bloch DA *et al*. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988; 31:315–24.

2 Aletaha D, Neogi T, Silman AJ *et al*. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010;69:1580–8.

3 Humphreys JH, Verstappen SM, Hyrich KL *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 2012.

4 Britsemmer K, Ursum J, Gerritsen M *et al*. Validation of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: slight improvement over the 1987 ACR criteria. Ann Rheum Dis 2011;70:1468–70.

5 Cader MZ, Filer A, Hazlehurst J *et al*. Performance of the 2010 ACR/EULAR criteria for rheumatoid arthritis: comparison with 1987 ACR criteria in a very early synovitis cohort. Ann Rheum Dis 2011;70:949–55.

6 Krabben A, Huizinga TW, van der Helm-van Mil AH. Undifferentiated arthritis characteristics and outcomes when applying the 2010 and 1987 criteria for rheumatoid arthritis. Ann Rheum Dis 2012;71:238–41.

7 Symmons DP, Barrett EM, Bankhead CR *et al*. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. Br J Rheumatol 1994;33:735–9.

Rheumatology 2013;52:1142–1144
doi:10.1093/rheumatology/ket019
Advance Access publication 5 March 2013

Efficacy of ambrisentan in the treatment of digital ulcers in patients with systemic sclerosis: a preliminary study

SIR, SSc is characterized by fibrotic thickening of the skin and internal organs and widespread vascular damage [1]. The first vascular event is RP and digital ulcers (DUs) are frequently occurring in SSc evolving to fingertip necrosis, gangrene and amputation with high impact on quality of life [2]. An important pathogenic role is played by endothelin, a molecule with a potent vasoconstrictor agent which is elevated in SSc [3]. The endothelin receptor antagonists have given significant results in the therapy of pulmonary arterial hypertension (PAH), while the only randomized clinical trial in the treatment of DUs is limited to bosentan [4, 5].

In our study, we evaluated the efficacy and tolerability of ambrisentan in the treatment of DUs in patients who failed bosentan. The study and the off-label use of this drug has been approved by the hospital ethics committee (comitato etico interaziendale A.S.O. San Giovanni Battista, A.S.O. C.T.O./C.R.F./Maria Adelaide Di Torino) and by the hospital pharmaceutical commission after submission of documentation supporting the treatment. All patients were given a detailed explanation of the study and informed consent was obtained. Six patients, aged between 40 and 68 years (mean age 53.7 ± 11.3 years), with SSc according to ACR Criteria [6], presenting DUs with an onset of between 6 and 9 months were consecutively recruited. Four of the six patients had limited ACA-positive SSc and two of them had SCL-70 antibody-positive diffuse SSc. The mean duration of SSc was 7.1 years with an average duration of onset of RP of 12.3 years. None of the six patients was suffering from PAH. These patients had been treated with ACE inhibitors, prostanoids and bosentan (at least for 6 months) suspended due to unsatisfactory results or side effects. The recruited patients continued to receive i.v. therapy with prostanoids. After 1 month of wash-out after bosentan therapy, patients were treated with ambrisentan. The recruitment period and analysis of the study were conducted between September 2011 and May 2012. Ambrisentan was administered at 5 mg/day and safety follow-up was performed every 4 weeks for 6 months. Each patient was given a diary to report at each visit as follows: date of onset and duration of RP, Raynaud's Condition Score (RCS) and number of daily attacks, visual analogue scale (VAS) for pain (1–10), ulcer onset and location.

At baseline and at each visit, blood samples for routine analysis and tests [e.g. pregnancy test, scleroderma HAQ disability index (HAQ-DI)] were performed. Categorical variables were compared using independent samples *t*-test and skewed outcome measures using Mann–Whitney *U*-test. A *P*-value <0.05 was considered statistically significant.

In four patients, all ulcers healed completely, while in two patients only one DU each was still evident at the end of the treatment. At week 24, the number of RP attacks was significantly decreased (Δ -3.10, $P = 0.01$), RP duration was decreased without reaching statistical significance (Δ -14.5, $P = 0.077$), the RCS was instead significantly improved (Δ -2.3, $P = 0.03$) as well as pain VAS (Δ -3.5, $P = 0.02$). The HAQ-DI did not give meaningful results even if there was a trend for improvement. During the 24 weeks of treatment, no new ulcers were observed and a significant number of ulcers healed completely—at baseline there were 2.67 ± 0.82 and at week 24 there were 0.33 ± 0.52 ($P < 0.03$). Videocapillaroscopy and the evaluation of pulmonary function (FVC, DL_{CO} and 6MWT) and heart (echocardiography, NT-proBNP and troponin) did not change during the study. No liver increase was observed during treatment with ambrisentan (Table 1).

A prospective uncontrolled long-term treatment with bosentan showed a significant decrease in the number of DUs [7]. To date, there are only few data, limited to individual experiences, on ambrisentan in DUs treatment [8]. This is the first report on the effect of ambrisentan in a limited number of SSc patients without PAH who were previously unsuccessfully treated with bosentan. This study showed significant reduction in the total number of DUs, with no appearance of new lesions in a winter

5.1.1 Supplementary results

5.1.2 Differences between patients with missing and non-missing baseline data

Table 5.1.1 shows the characteristics of patients with and without missing data in the analyses in section 5.1. There were no striking differences between these two groups.

Table 5.1.1 Characteristics of patients with no missing data at baseline and those with missing baseline data

	Missing baseline *n=445*	Non-missing *n=1643*
Female	304(68)	1074(65)
Age onset	53(39-69)	55(43-68)
Symptom duration *(weeks)*	23(13-42)	26(14-48)
Deaths	122(27)	466(28)
Person-years follow up	5612	20700
Crude mortality rate *(per 1000 person-years follow up)*	21.7	22.5

Categorical variables are reported as n(%). Continuous variables are reported as median (interquartile range)

5.1.3 Checking the proportional hazards assumption in results paper 5.1

Figure 5.1.1 shows the plots of the observed and expected survival probabilities of the two groups of patients (those who fulfilled the 2010 criteria at baseline and those who did not). These plots overlap therefore suggesting the proportional hazards assumption had not been violated. The results of the *estat phtest* confirmed this, $\chi^2 = 0.59$, $p=0.44$.

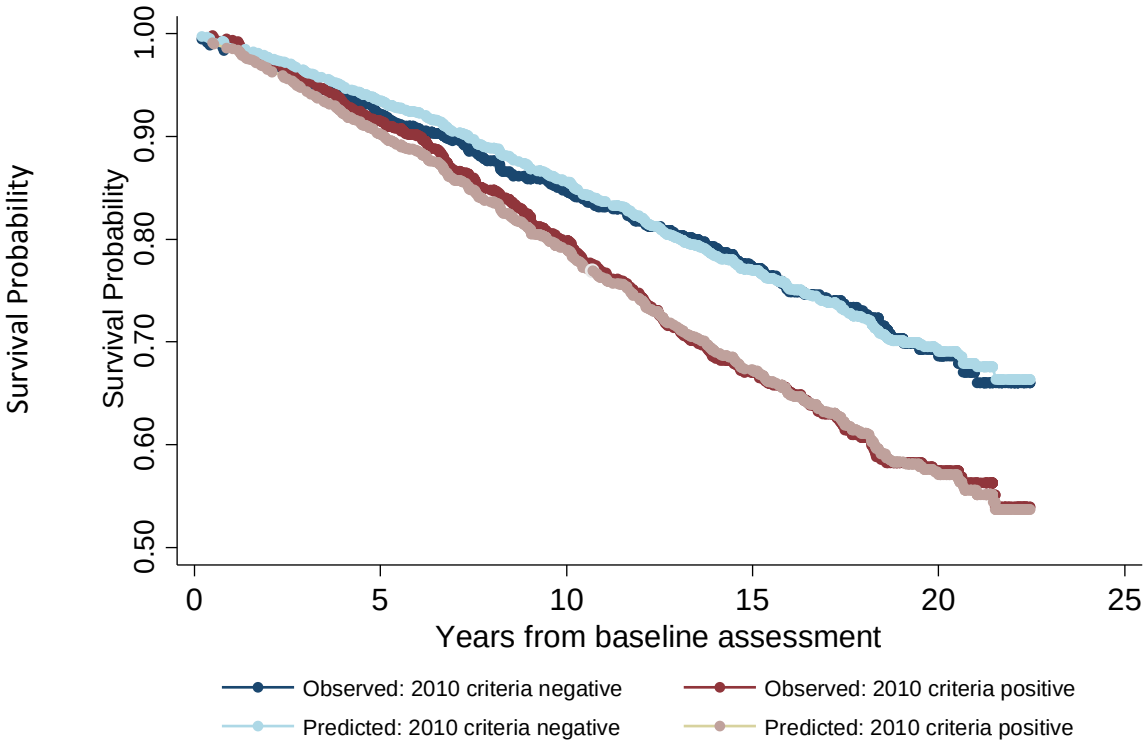


Figure 5.1.1 Observed and expected plots of survival probability in patients who do and do not fulfil the 2010 criteria at baseline

5.2 Rheumatoid Factor and Anti-Citrullinated Protein Antibody positivity, but not level, are associated with increased mortality in patients with rheumatoid arthritis: results from two large independent cohorts.

Humphreys JH, van Nies J, Chipping J, Marshall T, van der Helm-van Mil A, Symmons DP, Verstappen SM. *Arthritis Research & Therapy* 2014; 16:483

arthritis research&therapy

RESEARCH ARTICLE Open Access

Rheumatoid factor and anti-citrullinated protein antibody positivity, but not level, are associated with increased mortality in patients with rheumatoid arthritis: results from two large independent cohorts

Jennifer H Humphreys[1], Jessica AB van Nies[2], Jackie Chipping[3], Tarnya Marshall[4], Annette HM van der Helm-van Mil[2], Deborah PM Symmons[1,5] and Suzanne MM Verstappen[1*]

Abstract

Introduction: This study aimed to investigate rheumatoid factor (RF) and anti-citrullinated protein antibody (ACPA) status and levels as predictors of mortality in two large cohorts of patients with early inflammatory arthritis (EIA).

Methods: Data from the Norfolk Arthritis Register (NOAR) and Leiden Early Arthritis Clinic (EAC) cohorts were used. At baseline, patients had demographic data and smoking status recorded; RF, ACPA and inflammatory markers were measured in the local laboratories. Patients were flagged with national death registers until death or censor date. Antibody status was stratified as negative, low or high positive by RF and ACPA levels individually. In addition, patients were grouped as seronegative, RF positive, ACPA positive or double antibody (RF and ACPA) positive. Cox regression models explored associations between antibody status and mortality adjusting for age, sex, smoking status, inflammatory markers and year of enrolment.

Results: A total of 4962 patients were included, 64% were female. Median age at onset was 56 (NOAR) and 54 (EAC) years. In NOAR and EAC respectively, 35% and 42% of patients were ACPA/RF positive. When antibody status was stratified as negative, low or high positive, there were no consistent findings between the two cohorts. Double antibody positivity was associated with excess mortality in both cohorts compared to seronegative patients: NOAR and EAC respective adjusted HR (95% confidence interval) 1.35 (1.09 to 1.68) and 1.58 (1.16 to 2.15).

Conclusions: Patients with EIA who are seropositive for both RF and ACPA have increased mortality compared to those who are single positive or seronegative. Antibody level in seropositive patients was not consistently associated with excess mortality.

Introduction

In patients with inflammatory arthritis, the autoantibodies rheumatoid factor (RF) and anti-citrullinated protein antibody (ACPA) have been associated with poor outcomes, such as increased disease activity, radiographic progression and disability [1-5]. However, the utility of antibody level in predicting the prognosis of inflammatory arthritis, in particular rheumatoid arthritis (RA), has not been clearly established. In a recent multicentre prospective study of patients with early inflammatory arthritis (EIA), the presence of RF and/or ACPA was a significant predictor of RA diagnosis within two years, but level did not appear to be important [6]. In contrast, in a study of patients with EIA from Norway in 2010, Mjaavatten *et al.* found that increasing levels of RF and ACPA were associated with persistent joint inflammation [7]. Other studies have failed to show consistently that either RF or ACPA antibody level is

* Correspondence: suzanne.verstappen@manchester.ac.uk
[1]Arthritis Research UK Centre for Epidemiology, Manchester Academic Health Science Centre, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, UK
Full list of author information is available at the end of the article

important in predicting poor outcome in patients with EIA and RA [8-10]. In addition, recent data from a subset of the Leiden Early Arthritis Clinic have shown that the avidity of ACPA may be prognostically more important than the level itself [11].

Nevertheless, antibody level is included in the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria for RA [12], which aim to identify those patients with EIA with poor prognosis sufficient to require intervention with disease modifying therapy. The presence of RF and ACPA are weighted as part of the total score according to their level; patients are said to be low positive if their level is greater than the upper limit of normal (ULN) but less than three times the ULN, and high positive if their level is at least three times the ULN. Thus, patients with high antibody levels are more likely to fulfil the criteria, and it would be interesting to investigate whether these cut-offs are appropriate in predicting other adverse outcomes, such as mortality.

The increased mortality in patients with RA has been long established [13]. It is also well recognised that the presence of RF in sera of patients with inflammatory arthritis (whether or not they meet formal classification criteria for RA) is associated with an increased risk of premature death [14-16]. In fact, this association has been demonstrated even in subjects without symptoms of arthritis [17]. ACPA positivity has also been shown to predict premature mortality in the Norfolk Arthritis Register [18]; however this association has yet to be confirmed in other cohorts.

The aims of this study were to investigate the association between mortality and RF and/or ACPA positivity and level in patients with EIA. The term EIA includes all patients with RA early in the disease process, and studying these patients allows additional inclusion of those patients who may later go on to meet formal classification criteria for RA. It has been recognised that significant variability in antibody testing can occur between laboratories [19]. Thus, to strengthen the external validity of the study results, we investigated these questions in two large prospective cohorts of patients with EIA: the Norfolk Arthritis Register (NOAR) in the UK and the Leiden Early Arthritis Clinic (EAC) in the Netherlands.

Methods

Patients and setting

Patients in Norfolk, UK, were recruited to NOAR between 1990 and 2009 from primary and secondary care if they were adults (≥16 years) and had ≥2 swollen joints for ≥4 weeks; NOAR has been described in detail elsewhere [20]. Leiden EAC has also been described previously [21]; briefly patients in the region of Leiden, the Netherlands, with synovitis confirmed by a rheumatologist were recruited to the Leiden EAC from 1993 onwards if their symptom duration was less than two years at presentation. In order to make the two cohorts as comparable as possible, patients in NOAR were only included in this study if they had symptom duration of less than two years at presentation.

Assessment and follow up

Patients in NOAR are assessed at baseline by a research nurse who administers a structured questionnaire, including demographic details as well as disease and smoking history (never, past, current), performs a 51 tender and swollen joint count and obtains a blood sample. Sera are stored frozen and tested for C-reactive protein (CRP) and RF (latex test, low positive cut-off 40 units/ml, high positive cut-off 120 units/ml); subsequently ACPA, as defined by anti-CCP2 antibodies, are tested for using the Axis-Shield, Dundee, UK Diastat Anti-CCP kit (low positive cut-off 5 units/ml, high positive cut-off 15 units/ml). The Leiden EAC initial assessment includes medical history, clinical examination and joint counts. Blood samples are taken and tested for erythrocyte sedimentation rate (ESR), RF (IgM-RF in-house ELISA, low positive cut-off 5 units/ml, high positive cut-off 15 unit/ml) and ACPA (AntiCCP-2, Euro-Diagnostica, Malmo, Sweden ImmunoscanRA Mark 2, low positive cut-off 25 units/ml, high positive cut-off 75 units/ml). All cut-offs used are those recommended by the relevant manufacturers. Patients in NOAR are flagged with the NHS Information Centre (NHS IC) from baseline. NHS IC provide copies of death certificates to NOAR with approximately six months lag in reporting. They also provide a date of 'embarkation' for patients who leave the UK. Mortality data on patients recruited to the EAC are tracked nationally using the civic registries (Gemeentelijke Basis Administratie) in the Netherlands. NOAR is approved by Norfolk and Norwich University Hospital Local Research Ethics Committee UK, and EAC was approved by the local medical ethics committee LUMC The Netherlands.

Statistical analysis

Antibody levels were divided into negative, low positive and high positive as defined by the 2010 classification criteria [12]. These cut-offs were selected to investigate the ability of this aspect of the criteria to predict mortality. NOAR patients were censored for analysis at date of death, date of embarkation or 30 June 2012, whichever came first. Leiden EAC patients were censored at date of death or 1 May 2012. Analyses were conducted separately in each cohort. Kaplan-Meier survival curves were used to compare survival univariately in patients grouped according to their antibody status. Cox proportional hazard models were used to investigate the association between antibody status, antibody level and subsequent mortality.

A number of different models were developed. Firstly, patients were categorised according to antibody status as negative, low positive or high positive, and two models were then developed considering RF and ACPA status separately. A third model investigated whether the presence (above the ULN) of both antibodies, rather than antibody level, was important in predicting mortality by categorising patients as seronegative, RF single antibody positive, ACPA single antibody positive and double antibody positive (that is, both RF and ACPA positive). Univariate models were constructed initially, then age and sex adjusted; finally a multivariate model was developed adjusting for age, gender, baseline smoking status (categorised as current, ever or never smokers), inflammatory marker (ESR in EAC or CRP in NOAR) level, and year of enrolment to the cohort as a proxy for changing treatment strategies over time. All analyses were repeated in the population of patients fulfilling the 2010 ACR/EULAR criteria for RA. We aimed to focus on the predictive properties of the antibodies specifically and were deliberately parsimonious with our variable selection in the multivariate model. Thus, if a variable was not considered a confounder *a priori*, that is, would not have associations with both antibody status and mortality, it was not included. Similarly, variables that might be on the causal pathway between antibody status and mortality (such as disease activity over time) were also not included, as the relationship between antibody status and disease activity can only occur in one direction.

In the model in which the presence of both antibodies was compared to single antibody positivity and seronegativity, only patients who had been tested for both antibodies were included. In NOAR, 2,195 (72%) patients had data on both antibodies; data were more complete for the EAC, where 1,663 (87%) had both antibodies measured. In NOAR, therefore, baseline characteristics of patients with and without complete antibody data were assessed for differences. In addition, in order to ensure that the reported results were representative, multiple imputation using chained equations was performed to impute the antibody status of those patients with missing data. A subsequent sensitivity analysis was performed using the imputed dataset and these results were compared with those from the complete case analysis. Data from NOAR were analysed using the Stata 11 software package (Stata, College Station, TX, USA), data from EAC were analysed using SPSS for Windows version 20.0 (SPSS, Chicago, IL, USA).

Results

A total of 4,962 patients with EIA were included in the study (3,053 from NOAR, 1,909 from Leiden EAC). The cohorts had similar age and gender distributions, 65% (1,970) female in NOAR, 63% (1,205) female in the EAC, respective median (interquartile range) age at symptom onset 56 (44 to 68) and age at inclusion 54 (42 to 67) years. In NOAR, 63% of patients fulfilled the 2010 ACR/EULAR classification criteria for RA, in the EAC this proportion was 57% of patients. Baseline characteristics of patients from the two cohorts are shown in Table 1. The mean (standard deviation) follow up in each study was 11.8 (5.8) years in NOAR and 8.5 (5.2) years in EAC. There were 787 deaths during 36,109 person years follow up in NOAR, and 275 deaths during 16,187 person years follow up in the EAC; this resulted in crude death rates of 21.8 and 17.0 deaths per 1,000 person years in each cohort, respectively. The number of deaths in each of the antibody subgroups are shown in Table 2.

Table 1 Demographic and baseline disease characteristics

Demographic/characteristic	NOAR number = 3,053	Leiden EAC number = 1,909
Female *number (%)*	1970 (65)	1205 (63)
Age at symptom onset (years) *median (IQR)*	56 (44 to 68)	54 (42 to 67)
Symptom duration (weeks) *median (IQR)*	26 (14 to 47)	17 (8 to 33)
RF/ACPA positive umbern *(%)*	1079 (35)	810 (42)
RF positive	912 (34)	704 (37)
RF low positive	315 (12)	256 (13)
RF high positive	594 (22)	445 (23)
ACPA positive	598 (27)	591 (31)
ACPA low positive	91 (4)	66 (3.5)
ACPA high positive	507 (23)	532 (27.9)
Inflammatory marker (CRP, mg/L)	9 (2 to 20)	-
Inflammatory marker (ESR, mm/hr)	-	25 (11 to 44)
Smoking status		
Never	998 (33)	740 (45)
Previous	1189 (39)	445 (27)
Current	748 (26)	450 (28)
2010 ACR/EULAR RA criteria positive *number (%)*	1701 (63)	1073 (57)
1987 ACR RA criteria positive *number (%)*	1303 (43)	736 (39)

Categorical variables are presented as number (% non-missing data). % percentage missing values for NOAR and Leiden EAC, respectively, were as follows; RF/ACPA 10% and 0.5%, RF 11 % and 1%, ACPA 27% and 12%, CRP 18%, ESR 1%, smoking status 4% and 14%, 2010 RA 11% and 1%, 1987 RA 0.5% and 0%. ACPA, anti-citrullinated protein antibodies; ACR, American College of Rheumatology; CRP, C-reactive protein; EAC, Early Arthritis Clinic; ESR, erythrocyte sedimentation rate; EULAR European League Against Rheumatism; IQR, inter-quartile range; NOAR, Norfolk Arthritis Register; RF, rheumatoid factor; RA, rheumatoid arthritis. The italicised words describe how each characteristic is being presented numerically rather than the name of the characteristic itself, and are therefore italicised for clarity to make that distinction.

Table 2 Number of deaths in each antibody group

Antibody group	NOAR	Leiden EAC
RF/ACPA negative	401	28
RF/ACPA low positive	39	40
RF/ACPA high positive	264	106
RF negative	444	137
RF low positive	52	54
RF high positive	202	82
ACPA negative	394	154
ACPA low positive	21	17
ACPA high positive	156	86
Both antibodies negative	339	119
RF positive[a]	47	35
ACPA positive[a]	51	9
Both antibodies positive	128	93

[a]Where patients had both antibodies tested. ACPA, anti-citrullinated peptide antibodies; EAC, Early Arthritis Clinic; NOAR, Norfolk Arthritis Register; RF, rheumatoid factor.

Antibody levels

The first Cox proportional hazards models (univariate and adjusted) examined RF and ACPA levels separately (Table 3). There appeared to be a marked difference in RF high and low positivity in the NOAR cohort: low positive RF adjusted hazard ratio (HR) (95% confidence interval (CI)) 0.80 (0.59 to 1.08), high positive RF adjusted HR (95% CI) 1.49 (1.25 to 1.77). However, this was not replicated in the EAC cohort: low positive RF adjusted HR (95% CI) 1.62 (1.16 to 2.26), high positive RF adjusted HR (95% CI) 1.63 (1.19 to 2.24). Differences between the two cohorts were also seen with ACPA (Table 3). In the EAC, low positive ACPA status was associated with increased mortality, but high positive ACPA was not, respective adjusted HR (95% CI) 2.21 (1.31 to 3.72) and 1.25 (0.93 to 1.69). Conversely, in NOAR there was a trend towards increased mortality in the low positive ACPA group, and high positive ACPA status was significantly associated, adjusted HR (95% CI) 1.32 (1.08 to 1.61). Of note, there were only a small number of patients and, therefore, deaths in the ACPA low positive group in either cohort: 21 deaths in NOAR and 17 in the EAC. Similar findings were observed in the population of patients fulfilling the 2010 ACR/EULAR criteria for RA, although not always reaching statistical significance, probably due to smaller group sizes. Data on the full multivariate models are available as part of Additional file 1. The Additional file 1 also includes a model comparing patients negative for both antibodies to those with low and high levels of either antibody and models dividing RF and ACPA levels into tertiles rather than using the predefined cut-offs of the 2010 criteria. These additional models demonstrated similar results to those reported here.

Number of antibodies

This Cox model stratified patients by the number of antibodies present (negative, RF positive, ACPA positive, and double antibody positive if both RF and ACPA were positive). The results were more consistent between the two cohorts (Table 4 and Figure 1) and between the total EIA population and the 2010 ACR/EULAR RA population. In both NOAR and the EAC there was a trend towards increased mortality in patients who had a single positive antibody compared to no positive antibodies, other than single ACPA positivity in the Leiden EAC, where the number of deaths was small. In both cohorts the presence of two positive antibodies was significantly associated with increased mortality, adjusted HRs (95% CI) NOAR: 1.35 (1.09 to 1.68), EAC: 1.57 (1.15 to 2.14). No differences were identified in the baseline characteristics of patients with missing data in NOAR for this model, and the sensitivity analysis using imputed data produced similar results to the complete case analysis [see Additional file 1].

Discussion

In two well established observational cohorts of EIA and its sub-population of patients with RA, we have shown that RF and ACPA positivity are predictors of excess mortality, and that the presence of both antibodies was a stronger predictor of mortality than single antibody positivity. However, in this first large study to investigate the association between antibody levels and mortality, the influence of increasing antibody level was not consistent between the two cohorts.

Our data have once again demonstrated the known relationship between RF positivity and early mortality [14], and confirmed that a similar association exists in patients who are ACPA positive. This has previously been described in NOAR [18] but only reported elsewhere by two other groups of investigators. The first study was in a subset of 299 patients in the Rochester epidemiology project [22], half of whom had RA. The second small study, by Sihvonen *et al.* [23] used logistic regression (which does not allow for censoring) rather than Cox models to analyse the data. It was, therefore, important to corroborate this association in another large EIA cohort, such as the Leiden EAC.

The results of our study are concordant with the findings of Ursum *et al.*, who studied 545 patients with early arthritis in the Netherlands [10]. They found no association after two years between antibody levels and early disease outcomes, including disease activity measured by DAS28, functional status measured by the Health Assessment Questionnaire (HAQ) and radiographic progression. Similarly, a number of other small studies have reinforced the association between ACPA positivity and other poor outcomes, such as increased disease activity and

Table 3 Comparison of patients RF or ACPA negative to those with low and high RF or ACPA levels

	NOAR				Leiden EAC			
	Total EIA population		2010 ACR/EULAR positive cohort		Total EIA population		2010 ACR/EULAR positive cohort	
Model/predictor	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
RF:								
Unadjusted[a]								
RF low positive	0.90	0.68 to 1.20	0.90	0.66 to 1.24	2.13	1.55 to 2.91	1.75	1.21 to 2.54
RF high positive	1.67	1.41 to 1.97	1.43	1.18 to 1.74	1.75	1.33 to 2.31	1.39	0.99 to 1.93
Age and sex adjusted[a]								
RF low positive	0.81	0.61 to 1.08	0.80	0.58 to 1.10	1.67	1.21 to 2.29	1.67	1.15 to 2.42
RF high positive	1.54	1.30 to 1.82	1.33	1.09 to 1.62	1.92	1.46 to 2.53	2.00	1.42 to 2.81
Multivariate[ab]								
RF low positive	0.80	0.59 to 1.08	0.85	0.61 to 1.18	1.62	1.16 to 2.26	1.57	1.07 to 2.32
RF high positive	1.49	1.25 to 1.77	1.40	1.14 to 1.71	1.63	1.19 to 2.24	1.68	1.16 to 2.44
ACPA:								
Unadjusted[c]								
ACPA low positive	1.05	0.68 to 1.63	0.98	0.61 to 1.59	1.65	1.00 to 2.72	0.97	0.54 to 1.73
ACPA high positive	1.49	1.24 to 1.79	1.27	1.03 to 1.57	1.17	0.90 to 1.52	0.79	0.58 to 1.06
Age and sex adjusted[c]								
ACPA low positive	1.16	0.75 to 1.81	1.19	0.73 to 1.93	2.52	1.52 to 4.18	1.99	1.10 to 3.61
ACPA high positive	1.41	1.17 to 1.69	1.29	1.04 to 1.59	1.45	1.11 to 1.90	1.37	1.00 to 1.89
Multivariate[cb]								
ACPA low positive	1.39	0.89 to 2.16	1.44	0.89 to 2.36	2.21	1.31 to 3.72	1.78	0.96 to 3.28
ACPA high positive	1.32	1.08 to 1.61	1.24	0.99 to 1.57	1.25	0.93 to 1.69	1.22	0.86 to 1.73

[a]RF negative was used as a reference group; [b]adjusted for age at symptom onset, sex, baseline smoking status, year of inclusion in cohort and inflammatory marker; [c]ACPA negative was used as a reference group. Inflammatory marker = C-reactive protein in NOAR, = erythrocyte sedimentation rate in EAC. ACPA, anti-citrullinated protein antibodies; CI, confidence interval; HR, hazard ratio; RF, rheumatoid factor. The bold and italic text indicates subtitles, hence why there are no values in the table next to them. It is therefore essential that they look different to the predictor variables and the result values themselves.

Table 4 RF and ACPA positive versus single positive and both antibodies negative

	NOAR				Leiden EAC			
	Total EIA population		2010 ACR/EULAR positive cohort		Total EIA population		2010 ACR/EULAR positive cohort	
Model/predictor	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Unadjusted								
RF positive	1.11	0.83 to 1.49	1.10	0.79 to 1.53	1.88	1.29 to 2.74	1.53	0.99 to 2.36
ACPA positive	1.27	0.94 to 1.73	1.14	0.82 to 1.59	0.63	0.32 to 1.23	0.31	0.13 to 0.73
Both antibodies positive	1.51	1.23 to 1.85	1.29	1.02 to 1.64	1.59	1.21 to 2.09	1.12	0.80 to 1.57
Age and sex adjusted								
RF positive	1.05	0.78 to 1.41	1.10	0.79 to 1.54	1.45	0.99 to 2.13	1.54	0.99 to 2.37
ACPA positive	1.40	1.03 to 1.91	1.42	1.02 to 1.99	0.96	0.48 to 1.90	0.71	0.30 to 1.68
Both antibodies positive	1.38	1.12 to 1.69	1.25	0.99 to 1.59	1.82	1.38 to 2.40	1.83	1.29 to 2.60
Multivariate[a]								
RF positive	1.11	0.82 to 1.51	1.22	0.87 to 1.72	1.48	0.99 to 2.21	1.47	0.94 to 2.30
ACPA positive	1.35	0.98 to 1.88	1.39	0.97 to 1.99	1.05	0.53 to 2.09	0.79	0.33 to 1.89
Both antibodies positive	1.35	1.09 to 1.68	1.31	1.01 to 1.69	1.57	1.15 to 2.14	1.59	1.08 to 2.32

[a]Adjusted for age at symptom onset, sex, baseline smoking status, year of inclusion in cohort and inflammatory marker; both antibodies negative was used as reference group; inflammatory marker = C-reactive protein in NOAR, = erythrocyte sedimentation rate in EAC. ACPA, anti-citrullinated protein antibodies; CI, confidence interval; HR, hazard ratio; RF, rheumatoid factor. The bold and italic text indicates subtitles, hence why there are no values in the table next to them. It is therefore essential that they look different to the predictor variables and the result values themselves.

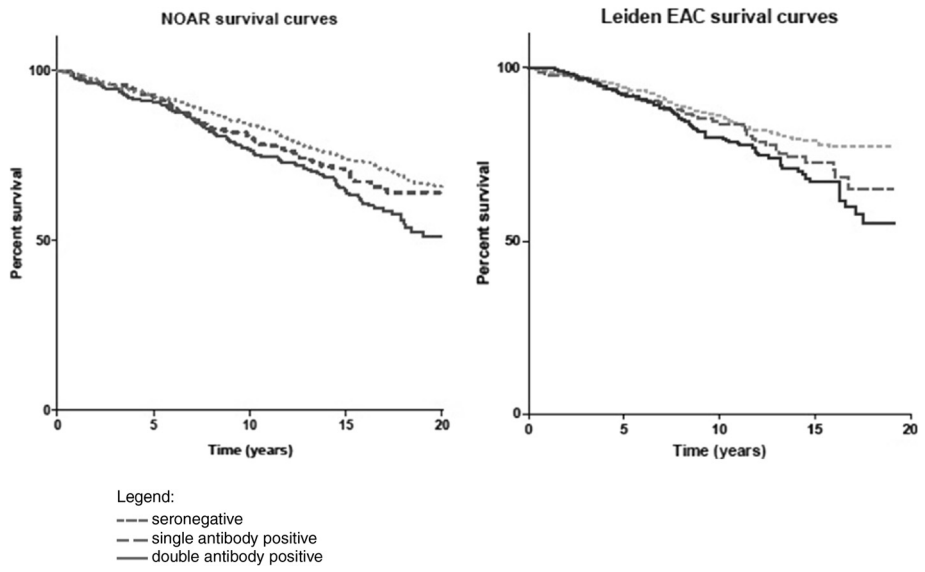


Figure 1 Unadjusted survival curves stratified by number of antibodies.

radiographic damage, but have failed to identify an association with increasing ACPA levels [8,24]. By contrast, Syversen *et al*. conducted a study of 125 patients who met the 1987 ACR classification criteria for RA [25] in a subpopulation of the European Research on Incapacitating Disease and Social Support (EURODISS) project [26]. They found that 10 year radiographic progression was increased in patients with low-moderate ACPA levels (>ULN and ≤8 times ULN), but this appeared to be further increased in patients with very high levels of ACPA (>8 times the ULN). However, they also demonstrated that the highest probability of radiographic progression occurred in patients who were positive for both RF and ACPA. A recent study in Italy examined progression from EIA to RA in 192 patients [6]. In accordance with our findings, they demonstrated the presence of both antibodies predicted RA, but antibody high or low positivity had no influence. In the Norwegian Very Early Arthritis Clinic (NOR-VEAC) study, Mjaavatten *et al*. showed additive value in testing for both antibodies in order to predict disease persistence [7]. They also demonstrated an association between antibody level and persistent arthritis, however the number of patients per group was small (<30). In addition, their analysis employed last observation carried forward to account for patients who did not have complete follow up. It is possible, therefore, that their results were influenced by attrition bias; that is, patients whose arthritis resolved may not have attended further follow up, and at their last recorded visit, their arthritis appeared to be persistent even though it subsequently resolved. It is possible that the different characteristics and follow up of these cohorts account for the different findings; in addition the different cut-offs of the commercially available assays may not correspond. Nevertheless, this emphasises that the role of antibody levels in predicting outcomes for patients with inflammatory arthritis has not been robustly established.

There are limitations to our study. We decided not to perform a pooled analysis of data from both cohorts because the different inclusion criteria of the two cohorts could potentially produce misleading conclusions. We did not aim to develop a full predictive model for mortality in RA, but focussed specifically on the association between antibody status and level, and mortality. Therefore, the number of confounders included in the multivariate model was small, and the final model does not account for all predictors of mortality in RA. As in all observational studies, there remains potential for residual confounding for which we have not adjusted. Further, in our analyses we did make the assumption that antibody status is fixed. This assumption seemed reasonable as the majority of studies have shown for both RF and, particularly, ACPA, that few patients convert from seropositive to negative over time [27-29], and when this does occur, risk of poor outcome may be maintained [30].

Conclusions

In conclusion, in this large study investigating the relationship between antibody levels and mortality in EIA, we have shown that patients with both RF and ACPA, rather than the higher levels of the antibodies, had increased rates of early death. We have also confirmed the association between ACPA positivity and excess mortality in a second large EIA cohort. Therefore, in patients presenting with early rheumatoid arthritis, the number

of positive antibodies may be more important than the antibody levels in assessing the mortality risk in clinical practice.

Additional file

Additional file 1: Table S1. Univariate and multivariate Cox proportional hazard models comparing RF/ACPA high/low positive versus negative. **Table S2.** Univariate and multivariate Cox proportional hazard models comparing RF and ACPA high/low positive versus negative. **Table S3.** Comparison of RF or ACPA levels in tertiles. **Table S4.** Univariate and multivariate Cox proportional hazard models comparing RFand ACPA positive versus single positive and both antibodies negative. **Table S5.** Sensitivity analysis with imputed data.

Abbreviations

ACPA: anti-citrullinated protein antibody; ACR: American College of Rheumatology; DAS28: Disease Activity Score based on 28 joint count; CRP: C-reactive protein; EAC: Leiden Early Arthritis Clinic; EIA: early inflammatory arthritis; ELISA: enzyme-linked immunosorbent assay; ESR: erythrocyte sedimentation rate; EULAR: European League Against Rheumatism; EURODISS: European Research on Incapacitating Disease and Social Support; HR: hazards ratio; NHS-IC: NHS Information Centre; IQR: inter-quartile range; NOAR: Norfolk Arthritis Register; RA: rheumatoid arthritis; RF: rheumatoid factor; ULN: upper limit of normal; 95% CI: 95% confidence interval.

Competing interests

All authors declare that they have no competing interests.

Authors' contributions

JH participated in conception and design of the study, data analysis of NOAR data, manuscript writing and final approval of the manuscript. JN participated in conception and design of the study, data analysis of Leiden data, help drafting and final approval of the manuscript. JC participated in data collection, drafting of manuscript and final approval of the manuscript. TM participated in data collection, drafting of manuscript and final approval of the manuscript. AH participated in conception and design of the study, revising the manuscript and final approval of the manuscript. DS participated in conception and design of the study, revising the manuscript and final approval of the manuscript. SV participated in conception and design of the study, revising the manuscript and final approval of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to acknowledge 'BeTheCure' for its support of collaborative work between European cohort studies. They also gratefully acknowledge the support of clinical staff in Norfolk, the NOAR research nurses, as well as the Leiden EAC clinical and research staff. NOAR is funded by Arthritis Research UK grant reference numbers: 20380 and 20385. JH Humphreys is funded by Arthritis Research UK grant reference number: 19743. This research has also been funded by The European Community Seventh Framework Program FP7 Health-F2-2008-223404 (Masterswitch) and a Vidi grant of the Netherlands Organization for Scientific Research.

Author details

[1]Arthritis Research UK Centre for Epidemiology, Manchester Academic Health Science Centre, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, UK. [2]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands. [3]Norfolk Arthritis Register, School of Medicine Health Policy and Practice, University of East Anglia, Norwich, UK. [4]Norfolk and Norwich University Hospital, Norwich, UK. [5]NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK.

Received: 28 March 2014 Accepted: 5 November 2014
Published online: 04 December 2014

References

1. Quinn MA, Gough AK, Green MJ, Devlin J, Hensor EM, Greenstein A, Fraser A, Emery P: **Anti-CCP antibodies measured at disease onset help identify seronegative rheumatoid arthritis and predict radiological and functional outcome.** *Rheumatology* 2006, **45:**478–480.
2. De Rycke L, Peene I, Hoffman IE, Kruithof E, Union A, Meheus L, Lebeer K, Wyns B, Vincent C, Mielants H, Boullart L, Serre G, Veys EM, De Keyser F: **Rheumatoid factor and anticitrullinated protein antibodies in rheumatoid arthritis: diagnostic value, associations with radiological progression rate, and extra-articular manifestations.** *Ann Rheum Dis* 2004, **63:**1587–1593.
3. Ronnelid J, Wick MC, Lampa J, Lindblad S, Nordmark B, Klareskog L, van Vollenhoven RF: **Longitudinal analysis of citrullinated protein/peptide antibodies (anti-CP) during 5 year follow up in early rheumatoid arthritis: anti-CP status predicts worse disease activity and greater radiological progression.** *Ann Rheum Dis* 2005, **64:**1744–1749.
4. Kudo-Tanaka E, Ohshima S, Ishii M, Mima T, Matsushita M, Azuma N, Harada Y, Katada Y, Ikeue H, Umeshita-Sasai M, Miyatake K, Saeki Y: **Autoantibodies to cyclic citrullinated peptide 2 (CCP2) are superior to other potential diagnostic biomarkers for predicting rheumatoid arthritis in early undifferentiated arthritis.** *Clin Rheumatol* 2007, **26:**1627–1633.
5. Bos WH, Bartelds GM, Vis M, van der Horst AR, Wolbink GJ, Van De Stadt RJ, Van Schaardenburg D, Dijkmans BA, Lems WF, Nurmohamed MT, Aarden L, Hamann D: **Preferential decrease in IgG4 anti-citrullinated protein antibodies during treatment with tumour necrosis factor blocking agents in patients with rheumatoid arthritis.** *Ann Rheum Dis* 2009, **68:**558–563.
6. Bizzaro N, Bartoloni E, Morozzi G, Manganelli S, Riccieri V, Sabatini P, Filippini M, Tampoia M, Afeltra A, Sebastiani G, Alpini C, Bini V, Bistoni O, Alunno A, Gerli R: **Anti-cyclic citrullinated peptide antibody titer predicts time to rheumatoid arthritis onset in patients with undifferentiated arthritis: results from a 2-year prospective study.** *Arthritis Res Ther* 2013, **15:**R16.
7. Mjaavatten MD, van der Heijde D, Uhlig T, Haugen AJ, Nygaard H, Sidenvall G, Helgetveit K, Kvien TK: **The likelihood of persistent arthritis increases with the level of anti-citrullinated peptide antibody and immunoglobulin M rheumatoid factor: a longitudinal study of 376 patients with very early undifferentiated arthritis.** *Arthritis Res Ther* 2010, **12:**R76.
8. Papadopoulos NG, Tsiaousis GZ, Pavlitou-Tsiontsi A, Giannakou A, Galanopoulou VK: **Does the presence of anti-CCP autoantibodies and their serum levels influence the severity and activity in rheumatoid arthritis patients?** *Clin Rev Allergy Immunol* 2008, **34:**11–15.
9. Shiozawa K, Kawasaki Y, Yamane T, Yoshihara R, Tanaka Y, Uto K, Shiozawa S: **Anticitrullinated protein antibody, but not its titer, is a predictor of radiographic progression and disease activity in rheumatoid arthritis.** *J Rheumatol* 2012, **39:**694–700.
10. Ursum J, Bos WH, van Dillen N, Dijkmans BA, Van Schaardenburg D: **Levels of anti-citrullinated protein antibodies and IgM rheumatoid factor are not associated with outcome in early arthritis patients: a cohort study.** *Arthritis Res Ther* 2010, **12:**R8.
11. Suwannalai P, Britsemmer K, Knevel R, Scherer HU, Levarht EW, van der Helm-van Mil AH, Van Schaardenburg D, Huizinga TW, Toes RE, Trouw LA: **Low-avidity anticitrullinated protein antibodies (ACPA) are associated with a higher rate of joint destruction in rheumatoid arthritis.** *Ann Rheum Dis* 2014, **73:**720–726.
12. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, Birnbaum NS, Burmester GR, Bykerk VP, Cohen MD, Combe B, Costenbader KH, Dougados M, Emery P, Ferraccioli G, Hazes JM, Hobbs K, Huizinga TW, Kavanaugh A, Kay J, Kvien TK, Laing T, Mease P, Menard HA, Moreland LW, Naden RL, Pincus T, Smolen JS, Stanislawska-Biernat E, Symmons D, *et al*: **2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative.** *Ann Rheum Dis* 2010, **69:**1580–1588.
13. Sokka T, Abelson B, Pincus T: **Mortality in rheumatoid arthritis: 2008 update.** *Clin Exp Rheumatol* 2008, **26:**S35–S61.
14. Goodson NJ, Wiles NJ, Lunt M, Barrett EM, Silman AJ, Symmons DP: **Mortality in early inflammatory polyarthritis: cardiovascular mortality is increased in seropositive patients.** *Arthritis Rheum* 2002, **46:**2010–2019.
15. Sokka T, Hakkinen A, Krishnan E, Hannonen P: **Similar prediction of mortality by the health assessment questionnaire in patients with rheumatoid arthritis and the general population.** *Ann Rheum Dis* 2004, **63:**494–497.

16. Solomon DH, Karlson EW, Rimm EB, Cannuscio CC, Mandl LA, Manson JE, Stampfer MJ, Curhan GC: **Cardiovascular morbidity and mortality in women diagnosed with rheumatoid arthritis.** *Circulation* 2003, **107:**1303–1307.
17. Tomasson G, Aspelund T, Jonsson T, Valdimarsson H, Felson DT, Gudnason V: **Effect of rheumatoid factor on mortality and coronary heart disease.** *Ann Rheum Dis* 2010, **69:**1649–1654.
18. Farragher TM, Goodson NJ, Naseem H, Silman AJ, Thomson W, Symmons D, Barton A: **Association of the HLA-DRB1 gene with premature death, particularly from cardiovascular disease, in patients with rheumatoid arthritis and inflammatory polyarthritis.** *Arthritis Rheum* 2008, **58:**359–369.
19. van der Linden MP, Batstra MR, Bakker-Jonges LE, Detert J, Bastian H, Scherer HU, Toes RE, Burmester GR, Mjaavatten MD, Kvien TK, Huizinga TW, van der Helm-van Mil AH: **Toward a data-driven evaluation of the 2010 American College of Rheumatology/European League against Rheumatism criteria for rheumatoid arthritis: is it sensible to look at levels of rheumatoid factor?** *Arthritis Rheum* 2011, **63:**1190–1199.
20. Symmons DP, Barrett EM, Bankhead CR, Scott DG, Silman AJ: **The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register.** *Br J Rheumatol* 1994, **33:**735–739.
21. de Rooy DP, van der Linden MP, Knevel R, Huizinga TW, van der Helm-van Mil AH: **Predicting arthritis outcomes–what can be learned from the Leiden Early Arthritis Clinic?** *Rheumatology (Oxford)* 2011, **50:**93–100.
22. Liang KP, Kremers HM, Crowson CS, Snyder MR, Therneau TM, Roger VL, Gabriel SE: **Autoantibodies and the risk of cardiovascular events.** *J Rheumatol* 2009, **36:**2462–2469.
23. Sihvonen S, Korpela M, Mustila A, Mustonen J: **The predictive value of rheumatoid factor isotypes, anti-cyclic citrullinated peptide antibodies, and antineutrophil cytoplasmic antibodies for mortality in patients with rheumatoid arthritis.** *J Rheumatol* 2005, **32:**2089–2094.
24. Landmann T, Kehl G, Bergner R: **The continuous measurement of anti-CCP-antibodies does not help to evaluate the disease activity in anti-CCP-antibody-positive patients with rheumatoid arthritis.** *Clin Rheumatol* 2010, **29:**1449–1453.
25. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, Healey LA, Kaplan SR, Liang MH, Luthra HS: **The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis.** *Arthritis Rheum* 1988, **31:**315–324.
26. Syversen SW, Gaarder PI, Goll GL, Odegard S, Haavardsholm EA, Mowinckel P, van der Heijde D, Landewe R, Kvien TK: **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:**212–217.
27. Barra L, Pope J, Bessette L, Haraoui B, Bykerk V: **Lack of seroconversion of rheumatoid factor and anti-cyclic citrullinated peptide in patients with early inflammatory arthritis: a systematic literature review.** *Rheumatology (Oxford)* 2011, **50:**311–316.
28. Mjaavatten MD, van der Heijde DM, Uhlig T, Haugen AJ, Nygaard H, Bjorneboe O, Kvien TK: **Should anti-citrullinated protein antibody and rheumatoid factor status be reassessed during the first year of followup in recent-onset arthritis? A longitudinal study.** *J Rheumatol* 2011, **38:**2336–2341.
29. Burr ML, Viatte S, Bukhari M, Plant D, Symmons DP, Thomson W, Barton A: **Long-term stability of anti-cyclic citrullinated peptide antibody status in patients with early inflammatory polyarthritis.** *Arthritis Res Ther* 2012, **14:**R109.
30. Nell-Duxneuner V, Machold K, Stamm T, Eberl G, Heinzl H, Hoefler E, Smolen JS, Steiner G: **Autoantibody profiling in patients with very early rheumatoid arthritis: a follow-up study.** *Ann Rheum Dis* 2010, **69:**169–174.

doi:10.1186/s13075-014-0483-3
Cite this article as: Humphreys *et al.*: Rheumatoid factor and anti-citrullinated protein antibody positivity, but not level, are associated with increased mortality in patients with rheumatoid arthritis: results from two large independent cohorts. *Arthritis Research & Therapy* 2014 **16**:483.

5.2.1 Supplementary results

5.2.2 Differences between patients with missing and non-missing data

Only data from NOAR are reported in this section, as the analysis in the Leiden EAC was performed independently. Table 5.2.1 shows no major differences in the baseline characteristics between patients with and without missing blood samples. Note in the non-missing group there remained a small number of patients (85, 3%) for whom baseline smoking status was missing.

Table 5.2.1 Baseline characteristics of patients with and without missing blood samples

	Missing *n=574*	Non-missing *n=2479*
Female	390(66)	1590(64)
Age onset	54(40-69)	56(44-68)
Symptom duration *(weeks)*	27(15-50)	26(14-46)
Smoking status: *current*	153(28)	595(24)
previous	198(34)	991(40)
never	190(33)	808(33)
Mean follow up *(years)*	11.8	11.8

Unless stated, categorical variables are reported as n(%) and continuous variables are reported as median (interquartile range)

5.2.3 Checking the proportional hazards assumption

As above, proportional hazards assumption could only be checked formally in the NOAR cohort. There were four Cox models reported. An example plot of the survival probabilities of patients grouped by the number of antibodies present is shown in figure 5.2.1. This demonstrates the overlap of observed and expected curves in each group. Table 5.2.2 shows the results of *estat phtest* for each of the models.

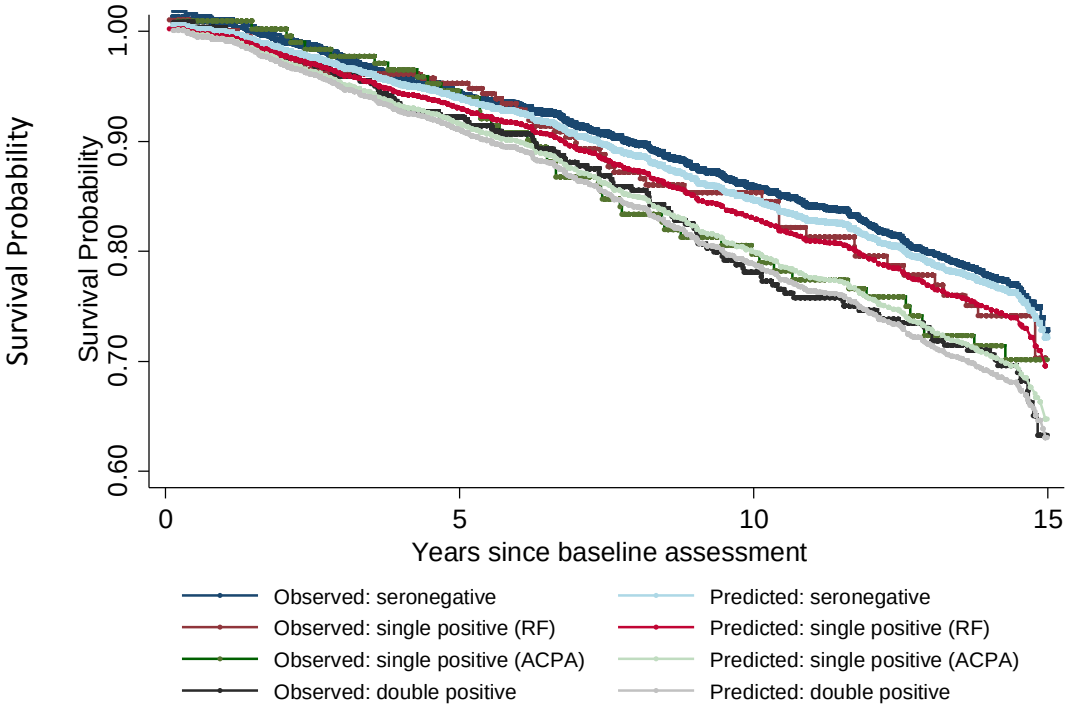


Figure 5.2.1 Observed and expected survival probabilities of patients grouped by number of antibodies

Table 5.2.2 Test of proportional hazards assumption for each model

Model	χ^2	P
ACPA/RF high/low positive vs negative	9.81	0.28
RF high/low positive vs negative	8.98	0.34
ACPA high/low positive vs negative	7.78	0.45
RF&ACPA positive vs single positive & double negative	4.38	0.88

ACPA: anti-citrullinated protein antibody, RF: rheumatoid factor

5.2.4 Online supplementary material for results paper 5.2

The following data in tables 5.2.3, 5.2.4 and 5.2.5 are referred to in the main manuscript in section 5.2 and formed the online supplementary material published to accompany this paper.

Table 5.2.3 Univariate and multivariate Cox proportional hazard models comparing RF/ACPA high/low positive vs negative (supplementary table 1)

	NOAR			Leiden EAC		
	Unadjusted	**Age & sex adjusted**	**Multivariate***	**Unadjusted**	**Age & sex adjusted**	**Multivariate***
	HR (95% CI)	*HR (95% CI)*	*HR (95% CI)*	*HR (95% CI)*	*HR (95% CI)*	*HR (95% CI)*
RF/ACPA negative	Ref	Ref	ref	Ref	ref	ref
RF/ACPA low positive	0.85(0.61-1.18)	0.81(0.59-1.10)	0.80(0.57-1.13)	2.38(1.67-3.39)	1.89(1.13-2.70)	1.92(1.30-2.80)
RF/ACPA high positive	1.58(1.35-1.84)	1.51(1.29-1.77)	1.44(1.21-1.70)	1.44(1.11-1.86)	1.63(1.26-2.12)	1.42(1.06-1.91)
Female gender	-	0.70(0.60-0.81)	0.77(0.65-0.91)	-	0.73(0.57-0.92)	0.72(0.54-0.94)
Age at onset	-	1.10(1.09-1.11)	1.11(1.10-1.12)	-	1.10(1.09-1.11)	1.10(1.09-1.12)
Smoking *Never*			ref			ref
Ever	-	-	0.98(0.81-1.18)	-	-	1.00(0.71-1.42)
Current			1.68(1.34-2.10)			1.81(1.32-2.49)
Inclusion year	-	-	0.96(0.94-0.98)	-	-	0.95(0.93-1.00)
Inflammatory marker	-	-	1.003(1.002-1.005)	-	-	1.006(1.001-1.011)

HR, hazard ratio; CI, confidence interval; RF, rheumatoid factor; ACPA, anti-citrullinated peptide antibodies; inflammatory marker = C-reactive protein in NOAR, = erythrocyte sedimentation rate in EAC

*adjusted for age at symptom onset, sex, baseline smoking status, year of inclusion in cohort & inflammatory marker

Table 5.2.4 Univariate and multivariate Cox proportional hazard models comparing RF and ACPA high / low positive vs negative (supplementary table 2)

		NOAR			Leiden EAC		
		Unadjusted	**Age & sex adjusted**	**Multivariate***	**Unadjusted**	**Age & sex adjusted**	**Multivariate***
		HR (95% CI)	*HR (95% CI)*	*HR (95% CI)*	*HR (95% CI)*	*HR (95% CI)*	*HR (95% CI)*
	RF negative	ref	ref	ref	ref	ref	ref
	RF low positive	0.90 (0.68-1.20)	0.81 (0.62-1.06)	0.80 (0.59-1.07)	2.13 (1.55-2.91)	1.67 (1.21-2.29)	1.62 (1.16-2.26)
RF	**RF high positive**	1.67 (1.41-1.97)	1.54 (1.29-1.83)	1.49 (1.23-1.80)	1.75 (1.33-2.31)	1.92 (1.46-2.53)	1.63 (1.19-2.24)
	Female gender	-	0.71 (0.62-0.83	0.78 (0.66-0.92)	-	0.75 (0.59-0.95)	0.73 (0.56-0.97)
	Age at onset	-	1.10 (1.09-1.11)	1.11 (1.10-1.12)	-	1.10 (1.09-1.11)	1.11 (1.09-1.12)
	Smoking *Never*	-	-	ref	-	-	ref
	Previous			0.98 (0.81-1.19)			0.99 (0.70-1.40)
	Current			1.66 (1.33-2.08)			1.76 (1.28-2.42)
	Inclusion year	-	-	0.96 (0.94-0.98)	-	-	0.95 (0.92-0.99)
	Inflammatory marker	-	-	1.003 (1.002-1.005)	-	-	1.006 (1.001-1.010)
	ACPA negative	ref	ref	ref	ref	ref	ref
	ACPA low positive	1.05 (0.69-1.60)	1.16 (0.80-1.68)	1.38 (0.98-1.97)	1.65 (1.00-2.72)	2.52 (1.52-4.18)	2.21 (1.31-3.72)
ACPA	**ACPA high positive**	1.49 (1.24-1.79)	1.41 (1.15-1.72)	1.38 (1.11-1.71)	1.17 (0.90-1.52)	1.45 (1.11-1.90)	1.25 (0.93-1.69)
	Female gender	-	0.71 (0.60-0.84)	0.81 (0.67-0.98)	-	0.71 (0.55-0.91)	0.73 (0.55-0.97)
	Age at onset	-	1.10 (1.09-1.11)	1.11 (1.10-1.12)	-	1.10 (1.09-1.11)	1.10 (1.09-1.12)
	Smoking *Never*			ref			ref
	Previous	-	-	1.00 (0.81-1.24)	-	-	1.04 (0.73-1.48)
	Current			1.80 (1.40-2.31)			1.86 (1.35-2.56)
	Inclusion year	-	-	0.96 (0.94-0.98)	-	-	0.97 (0.93-1.01)
	Inflammatory marker	-	-	1.003 (1.001-1.005)	-	-	1.006 (1.001-1.011)

HR, hazard ratio; CI, confidence interval; RF, rheumatoid factor; ACPA, anti-citrullinated peptide antibodies; inflammatory marker = C-reactive protein in NOAR, = erythrocyte sedimentation rate in EAC
*adjusted for age at symptom onset, sex, baseline smoking status, year of inclusion in cohort & inflammatory marker

Table 5.2.5 Comparison of RF or ACPA levels in tertiles (supplementary table 3)

	NOAR		Leiden EAC	
	Total EIA population	**2010 criteria positive**	**Total EIA population**	**2010 criteria positive**
RF:	*HR (95% CI)*	*HR (95%CI)*	*HR (95% CI)*	*HR (95%CI)*
Unadjusted [†]				
RF second tertile	0.92(0.69-1.23)	0.73(0.47-1.13)	1.37(0.96-1.94)	1.03(0.65-1.64)
RF third tertile	1.39 0.19-1.63)	1.21(1.00-1.47)	1.91(1.47-2.49)	1.40(1.00-1.96)
Age & sex adjusted [†]				
RF second tertile	0.83(0.62-1.10)	0.85(0.55-1.31)	1.49(1.05-2.12)	1.46(0.91-2.34)
RF third tertile	1.25(1.07-1.47)	1.14(0.95-1.38)	1.86(1.43-2.42)	1.90(1.35-2.67)
Multivariate [†] *****				
RF second tertile	1.16(0.85-1.59)	1.18(0.74-1.90)	1.53(1.05-2.22)	1.42(0.87-2.32)
RF third tertile	1.29(1.09-1.52)	1.25(1.03-1.53)	1.66(1.23-2.23)	1.68(1.17-2.43)
ACPA:				
Unadjusted [§]				
ACPA second tertile	1.04(0.85-1.27)	1.25(0.94-1.66)	1.22(0.89-1.68)	1.18(0.78-1.78)
ACPA third tertile	1.29(1.06-1.57)	1.31(1.03-1.68)	1.26(0.91-1.73)	0.82(0.56-1.19)
Age & sex adjusted [§]				
ACPA second tertile	0.95(0.77-1.16)	1.16(0.87-1.54)	1.58(1.15-2.18)	1.58(1.04-2.38)
ACPA third tertile	1.18(0.97-1.44)	1.26(0.99-1.61)	1.77(1.28-2.46)	1.59(1.08-2.34)
Multivariate [§] *****				
ACPA second tertile	1.00(0.81-1.25)	1.24(0.91-1.67)	1.21(0.84-1.74)	1.22(0.76-1.95)
ACPA third tertile	1.19(0.97-1.46)	1.27(0.98-1.63)	1.30(0.90-1.87)	1.26(0.81-1.95)

HR, hazard ratio; CI, confidence interval; RF, rheumatoid factor; ACPA, anti-citrullinated protein antibodies; inflammatory marker = C-reactive protein in NOAR, = erythrocyte sedimentation rate in EAC

[†]RF first tertile was used as a reference group; [§]ACPA first tertile was used as a reference group

*adjusted for age at symptom onset, sex, baseline smoking status, year of inclusion in cohort & inflammatory marker

Table 5.2.6 Univariate and multivariate Cox proportional hazard models comparing RF&ACPA positive vs single positive & both antibodies negative (supplementary table 4)

	NOAR			Leiden EAC		
	Unadjusted	**Age & sex adjusted**	**Multivariate***	**Unadjusted**	**Age & sex adjusted**	**Multivariate***
	HR (95% C)I	*HR (95% CI)*	*HR (95% CI)*	*HR (95% CI)*	*HR (95% CI)*	*HR (95% CI)*
Both antibodies negative	ref	ref	ref	Ref	ref	Ref
RF positive	1.11 (0.83-1.49)	1.05 (0.78-1.41)	1.11 (0.82-1.51)	1.88 (1.29-2.74)	1.45 (0.99-2.13)	1.48 (0.99-2.21)
ACPA positive	1.27 (0.94-1.73)	1.40 (1.03-1.91)	1.35 (0.98-1.88)	0.63 (0.32-1.23)	0.96 (0.48-1.90)	1.05 (0.53-2.09)
Both antibodies positive	1.51 (1.23-1.84)	1.38 (1.11-1.72)	1.35 (1.09-1.68)	1.59 (1.21-2.09)	1.82 (1.38-2.40)	1.57 (1.15-2.14)
Female gender	-	0.72 (0.61-0.85)	0.81 (0.67-0.98)	-	0.72 (0.57-0.93)	0.74 (0.55-0.98)
Age at onset	-	1.11 (1.10-1.12)	1.11 (1.10-1.12)	-	1.10 (1.09-1.11)	1.10 (1.09-1.11)
Smoking *Never*			ref			ref
Previous	-	-	1.00 (0.81-1.24)	-	-	1.00 (0.70-1.42
Current			1.76 (1.37-2.27)			1.82 (1.31-2.51)
Inclusion year	-	-	0.96 (0.94-0.98)	-	-	0.96 (0.93-1.00)
Inflammatory marker	-	-	1.003 (1.002-1.005)	-	-	1.005 (1.000-1.010)

HR, hazard ratio; CI, confidence interval, inflammatory marker = C-reactive protein in NOAR, = erythrocyte sedimentation rate in EAC
*adjusted for age at symptom onset, sex, baseline smoking status, year of inclusion in cohort & inflammatory marker

Table 5.2.7 Sensitivity analysis with imputed data (supplementary table 5)

	NOAR		
		Multivariate*	
	HR	*95% CI*	*FMI*
Both antibodies negative	ref	-	-
RF positive	1.12	0.83-1.52	0.153
ACPA positive	1.31	0.95-1.83	0.198
Both antibodies positive	1.38	1.13-1.68	0.101
Female gender	0.76	0.65-0.90	0.002
Age at onset	1.11	1.10-1.12	0.005
Smoking *Never*	ref	-	
Ever	0.97	0.80-1.18	0.001
Current	1.65	1.32-2.07	0.003
Inclusion year	0.96	0.94-0.98	0.002
Inflammatory marker	1.003	1.002-1.005	0.004

HR, hazard ratio; CI, confidence interval; FMI, fraction of missing information;
inflammatory marker = C-reactive protein in NOAR
*adjusted for age at symptom onset, sex, baseline smoking status, year of inclusion in cohort & inflammatory marker

5.3 Mortality trends in patients with early rheumatoid arthritis over 20 years: results from the Norfolk Arthritis Register.

Humphreys JH, Warner A, Chipping J, Marshall T, Lunt M, Symmons DP, Verstappen SM. *Arthritis Care & Research* 2014; 66(9):1296-301

Arthritis Care & Research
Vol. 66, No. 9, September 2014, pp 1296–1301
DOI 10.1002/acr.22296

ORIGINAL ARTICLE

Mortality Trends in Patients With Early Rheumatoid Arthritis Over 20 Years: Results From the Norfolk Arthritis Register

J. H. HUMPHREYS,[1] A. WARNER,[1] J. CHIPPING,[2] T. MARSHALL,[3] M. LUNT,[1] D. P. M. SYMMONS,[4] AND S. M. M. VERSTAPPEN[1]

Objective. **To examine mortality rates in UK patients with early rheumatoid arthritis (RA) from 1990–2011 and compare with population trends.**

Methods. **The Norfolk Arthritis Register (NOAR) recruited adults with ≥2 swollen joints for ≥4 weeks: cohort 1 (1990–1994), cohort 2 (1995–1999), and cohort 3 (2000–2004). At baseline, serum rheumatoid factor and anti–citrullinated protein antibody were measured and the 2010 American College of Rheumatology/European League Against Rheumatism RA classification criteria were applied. Patients were followed for 7 years, until emigration or death. The UK Office for National Statistics notified the NOAR of the date and cause of deaths, and provided mortality rates for the Norfolk population. All-cause and cardiovascular-specific standardized mortality ratios (SMRs) were calculated. Poisson regression was used to compare mortality rate ratios (MRRs) between cohorts and then, with cubic splines, to model rates by calendar year. Analyses were performed in patients 1) with early inflammatory arthritis, 2) classified as having RA, and 3) autoantibody positive.**

Results. **A total of 2,517 patients were included, with 1,639 women (65%) and median age 55 years, and 1,419 (56%) fulfilled the 2010 RA criteria. All-cause and cardiovascular-specific SMRs were significantly elevated in the antibody-positive groups. There was no change in mortality rates over time after accounting for changes in the population rates. In RA patients, all-cause MRRs, compared to cohort 1, were 1.13 (95% confidence interval [95% CI] 0.84–1.52) and 1.00 (95% CI 0.70–1.43) in cohorts 2 and 3, respectively.**

Conclusion. **Mortality rates were increased in patients with RA and SMRs were particularly elevated in those who were autoantibody positive. Compared to the general population, mortality rates have not improved over the past 20 years.**

INTRODUCTION

It is well recognized that patients with rheumatoid arthritis (RA) die prematurely (1). Meta-analysis of studies published over the last 50 years suggest the standardized mortality ratio (SMR) is 1.47 (95% confidence interval [95% CI] 1.19–1.83) (2), i.e., patients with RA have a 47% increased risk of death compared to the general population, matched for age and sex. Causes of death in RA popula-

The Norfolk Arthritis Register was supported by Arthritis Research UK (grant 20380). Dr. Humphreys's work was supported by Arthritis Research UK (grant 19743).

[1]J. H. Humphreys, MBChB, MRCP, A. Warner, MBChB, M. Lunt, PhD, S. M. M. Verstappen, PhD: Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Institute of Inflammation and Repair, The University of Manchester, Manchester Academic Health Science Centre, Manchester, UK; [2]J. Chipping, BSc: Norfolk Arthritis Register, School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, UK; [3]T. Marshall, MD, FRCP: Norfolk and Norwich University Hospital, Norwich, UK; [4]D. P. M. Symmons, MD, FRCP: Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Institute of Inflammation and Repair, The University of Manchester, Manchester Academic Health Science Centre, and NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust and University of Manchester partnership, Manchester, UK.

Address correspondence to S. M. M. Verstappen, PhD, Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Institute of Inflammation and Repair, The University of Manchester, Manchester Academic Health Science Centre, Manchester, UK. E-mail: suzanne.verstappen@manchester.ac.uk.

Submitted for publication July 26, 2013; accepted in revised form January 28, 2014.

Significance & Innovations

- All-cause and cardiovascular-specific mortality are increased in patients who satisfy the 2010 American College of Rheumatology/European League Against Rheumatism criteria for rheumatoid arthritis (RA) compared to the general population.
- All-cause and, in particular, cardiovascular-specific mortality in patients with early inflammatory arthritis in the first 7 years appear to be decreasing over time.
- The rate of decrease in mortality in patients with RA over the last 20 years is similar to the rate of decrease in mortality in the general population.

tions are similar to those in the wider population, with cardiovascular disease (CVD) being the most common cause (3). There is good evidence that improvements in CVD treatment, alongside public health interventions to aid primary and secondary prevention, have led to a fall in CVD mortality in the UK over the last 30 years (4). In addition, the prevalence of smoking, an important risk factor for both CVD and RA (5), has decreased in the UK by approximately 20% since 1980 (6). In RA, more aggressive treatment strategies and earlier intervention have also improved outcomes (7). Therefore, we might hypothesize that mortality in RA populations may also have improved, and may be approaching that of the general population. Indeed, some studies of patients with prevalent RA have suggested that such improvements have occurred (1). However, studies of prevalent cases are vulnerable to survivor bias, whereby a patient has to have survived with the disease long enough to be included in the study. By contrast, in a large incident cohort study from Rochester, Minnesota, Gonzalez et al reported that mortality rates increased over 40 years from 1965–2005 compared to the general population (8). They suggested that this was due to population-level improvements in mortality not being reflected in the RA population. Their study was limited to patients who fulfilled the 1987 American College of Rheumatology (ACR) classification criteria for RA (9) at baseline, which are poorly sensitive in early RA (10), and mortality is also recognized to be increased in patients with early inflammatory arthritis (EIA) before classification criteria have been met (11,12). Few large prospective cohort studies exist that are able to examine secular trends. In addition, we now have a new case definition of RA in the 2010 ACR/European League against Rheumatism (EULAR) classification criteria (13). The aim of this study was to describe trends in mortality, first among a cohort of patients with EIA, second in the subset of patients with RA defined by the 2010 RA criteria, and third in those positive for the autoantibodies rheumatoid factor (RF) and anti–citrullinated protein antibodies (ACPAs).

PATIENTS AND METHODS

Setting. This study was based in the Norfolk Arthritis Register (NOAR), UK. The NOAR has been described in detail elsewhere (14); briefly, since 1990 it has aimed to recruit adults registered with a general practitioner in the former Norwich Health Authority area presenting for the first time to primary or secondary care with EIA, defined as ≥2 swollen joints for ≥4 weeks. Patients recruited into the NOAR were divided into 3 cohorts, depending on the calendar year in which they were first enrolled on the register: cohort 1 (1990–1994), cohort 2 (1995–1999), or cohort 3 (2000–2004). All patients included in this study had <2 years' symptom duration at baseline assessment (89% of the total study population).

Assessment and followup. All patients recruited to the NOAR were seen by a research nurse at baseline, who conducted a structured interview and performed a 51 tender and swollen joint count. Blood samples were taken and the sera were stored frozen and later analyzed for C-reactive protein, RF (latex test), and ACPA (Axis-Shield Diastat anti-CCP kit). The 2010 ACR/EULAR criteria (13) were applied retrospectively using data collected at the baseline assessment. All patients were flagged with the Office for National Statistics (ONS), who notified dates of death to the NOAR and provided copies of death certificates. Deaths were attributed to CVD if the underlying cause of death was coded according to chapter I of the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (15). For any NOAR patients who left the UK, the ONS provided a date of embarkation; these patients were censored at that date. The ONS also provided age-, sex-, and cause-specific mortality rates by calendar year for the Norfolk County population, which covers a similar area to the patient population. The NOAR was approved by the Norwich Local Research Ethics Committee and all patients gave written consent.

Statistical analysis. Patients in each cohort were censored after 7 years of followup. This followup time was selected because it allowed inclusion of the most recent population mortality data available from the ONS (all deaths occurring prior to January 1, 2012) and provided a standardized length of time in which deaths could occur in each cohort, in order to facilitate comparisons. For each NOAR cohort, all-cause and CV-specific crude mortality rates were determined and 7-year SMRs were calculated by comparing the observed number of deaths to the expected number of deaths based on contemporary age- and sex-specific mortality rates from the ONS. This was done in the total population of EIA, for those classified as having RA according to the 2010 criteria, and in the subgroup of patients who were RF and/or ACPA positive. SMRs were not calculated if there were <15 observed deaths within a cohort, since CIs would be very wide and therefore it would not be possible to obtain a meaningful estimate. Mortality rate ratios (MRRs) were calculated using Poisson regression. MRRs allow statistical comparison of the mor-

Table 1. Demographic and baseline disease characteristics*

	Cohort 1	Cohort 2	Cohort 3	Total
EIA, n	1,010	879	628	2,517
Women	655 (65)	569 (65)	407 (65)	1,631 (65)
Age at symptom onset, median (IQR) years	54 (42–67)	55 (44–67)	58 (47–70)	55 (44–68)
Symptom duration, median (IQR) weeks	22 (12–41)	28 (16–51)	29 (17–49)	26 (14–46)
RF/ACPA positive	299 (34)	287 (36)	235 (42)	821 (37)†
2010 ACR/EULAR RA criteria positive	629 (69)	451 (57)	339 (61)	1,419 (63)‡
2010 ACR/EULAR RA criteria negative	287 (31)	337 (43)	218 (39)	842 (37)‡
1987 ACR RA criteria positive	458 (45)	318 (36)	289 (46)	1,065 (42)
DAS28, median (IQR)	3.97 (2.89–5.05)	3.54 (2.64–4.66)	3.60 (2.65–4.53)	3.71 (2.75–4.78)§
DMARDs at baseline assessment	153 (15)	258 (29)	287 (46)	698 (28)
2010 RA criteria positive, n	629	451	339	1,419
Women	412 (66)	313 (69)	234 (69)	959 (68)
Age at symptom onset, median (IQR) years	56 (44–68)	57 (47–69)	59 (49–69)	57 (47–68)
Symptom duration, median (IQR) weeks	23 (13–41)	28 (16–52)	31 (19–51)	26 (15–47)
RF/ACPA positive	270 (48)	247 (58)	198 (63)	715 (55)†
1987 ACR RA criteria positive	411 (65)	275 (61)	241 (71)	927 (65)
DAS28, median (IQR)	4.61 (3.85–5.58)	4.56 (3.55–5.37)	4.31 (3.54–4.98)	4.5 (3.68–5.40)§
DMARDs at baseline assessment	128 (20)	170 (38)	185 (55)	483 (34)

* Values are the number (% nonmissing data) unless indicated otherwise. EIA = early inflammatory arthritis; IQR = interquartile range; RF = rheumatoid factor; ACPA = anti–citrullinated protein antibody; ACR = American College of Rheumatology; EULAR = European League Against Rheumatism; RA = rheumatoid arthritis; DAS28 = 28-joint Disease Activity Score; DMARDs = disease-modifying antirheumatic drugs.
† Missing 280 (11%) EIA and 117 (8%) RA.
‡ Missing 257 (10%) positive and 0 (0%) negative.
§ Missing 458 (23%) EIA and 217 (15%) RA.

tality rates between the cohorts, while accounting for the expected mortality rate in each cohort based on age- and sex-specific mortality rates in the Norfolk population as an exposure variable. Cohort 1 was used as the reference standard. Mortality rates were then modeled by calendar year also using Poisson regression. A multivariate model, adjusted for age at symptom onset and sex and disease duration at baseline, and cubic splines were used to smooth the polynomial relationship between calendar year and mortality rates. All data were analyzed using the Stata 11 software package.

RESULTS

A total of 2,517 patients were included in this analysis, with 16,485 person-years of followup. A total of 1,419 patients (56%) fulfilled the 2010 RA criteria at baseline, 1,639 (65%) were women, and the median age at symptom onset was 55 years (interquartile range 44–68 years). Baseline demographic and clinical details for the 3 cohorts are shown in Table 1. The median age at onset increased with each succeeding cohort, as did the median symptom duration. Crude 7-year mortality rates generally decreased slightly over time: in cohorts 1, 2, and 3, they were 21.25, 21.43, and 19.96 per 1,000 person-years for all-cause mortality, respectively, and were 8.78, 7.87, and 7.07 per 1,000 person-years for CV-specific mortality, respectively.

Across the entire time span (all cohorts combined), the all-cause SMR was significantly elevated for the total EIA group (1.16, 95% CI 1.04–1.29) (Table 2). The SMR was higher for the patients fulfilling the RA criteria (1.22, 95% CI 1.07–1.40), but not for patients who did not fulfill the 2010 criteria for RA at baseline (0.90, 95% CI 0.73–1.11). The highest SMR was observed in the antibody-positive subgroup (1.39, 95% CI 1.18–1.65).

SMRs were calculated cross-sectionally for each of the 3 consecutive cohorts of patients recruited to the NOAR between 1990 and 2004. There did not appear to be any trend in the SMRs over time among the total EIA population, patients with RA, or those who were antibody positive (Table 2). This was confirmed in the Poisson regression, which assessed differences in mortality incidence after taking account of changes within the background population. This analysis showed no significant change in the MRRs in cohorts 2 and 3 compared to cohort 1 for patients with EIA, with RA, or who were antibody positive (Table 3).

CV mortality was not significantly elevated compared to the general population for the total EIA group overall or for any of the time cohorts. CV mortality was significantly elevated in the antibody-positive subgroup of cohort 1 (SMR 1.87, 95% CI 1.30–2.69). There were insufficient deaths in cohorts 2 and 3 to examine the CV-specific SMR. There was a nonsignificant trend toward increasing CV-specific MRR in the Poisson regression model for the RA subgroup. Again, there were insufficient numbers of deaths to explore this for the antibody subgroup (Tables 2 and 3).

Overall, persistently increasing, but stable, mortality rates over time were seen in patients with RA, when modeled by calendar year (Figures 1A and B). The Poisson regression used to create these plots demonstrated no evidence of change in all-cause or CV-specific mortality over time ($P = 0.92$ and 0.40, respectively).

Table 2. All-cause and cardiovascular-specific deaths and SMRs by cohort after 7 years of followup*

	No. of observed deaths				SMR (95% CI)			
	EIA	2010 RA criteria positive	2010 RA criteria negative	RF/ACPA positive	EIA	2010 RA criteria positive	2010 RA criteria negative	RF/ACPA positive
All cause								
Cohort 1	141	91	28	55	1.21 (1.02–1.41)	1.18 (0.96–1.45)	0.99 (0.69–1.44)	1.54 (1.18–2.00)
Cohort 2	123	75	36	44	1.17 (0.98–1.40)	1.30 (1.04–1.63)	0.89 (0.64–1.23)	1.25 (0.93–1.68)
Cohort 3	82	44	25	36	1.06 (0.85–1.32)	1.19 (0.89–1.60)	0.84 (0.57–1.24)	1.39 (1.00–1.92)
Total	346	210	89	135	1.16 (1.04–1.29)	1.22 (1.07–1.40)	0.90 (0.73–1.11)	1.39 (1.18–1.65)
Cardiovascular								
Cohort 1	58	36	9	29	1.16 (0.90–1.50)	1.11 (0.80–1.54)	–†	1.87 (1.30–2.69)
Cohort 2	45	26	15	12	1.07 (0.80–1.43)	1.13 (0.77–1.66)	0.92 (0.55–1.52)	–†
Cohort 3	29	15	9	10	1.02 (0.71–1.47)	1.19 (0.72–1.98)	–†	–†
Total	132	77	33	51	1.09 (0.92–1.30)	1.13 (0.91–1.42)	0.85 (0.60–1.19)	1.31 (1.00–1.73)

* SMR = standardized mortality ratio; 95% CI = 95% confidence interval; EIA = early inflammatory arthritis; RA = rheumatoid arthritis; RF = rheumatoid factor; ACPA = anti–citrullinated protein antibody.
† Too few events to calculate the SMR.

DISCUSSION

We have shown that all-cause mortality in the first 7 years of EIA and RA, defined according to the 2010 ACR/EULAR criteria, is higher than that in the general population, but the SMR has remained stable over the past 20 years. We have demonstrated, for the first time, increased SMRs in patients classified as having RA according to the 2010 ACR/EULAR classification criteria. In addition, crude all-cause and CV mortality rates in the first 7 years from baseline assessment in these patient groups decreased slightly over time; however, this decrease is occurring at the same rate as in the general population.

We were unable to confirm the findings of Gonzalez et al of a widening mortality gap between patients with RA and the general population (8). This may be because we have identified trends in mortality emerging in the 5 years since their study was completed in January 2007. Alternatively, it may be due to case definition; we used the 2010 classification criteria at baseline assessment to define RA, whereas in their study, incident cases of RA were recruited into the study when they fulfilled 4 of 7 of the 1987 criteria, which is likely to be further into the disease process than our baseline assessment. In addition, we restricted our analysis to deaths within the first 7 years of followup in order to standardize comparisons between the cohorts, whereas median followup in Minnesota was 11.7 years, which may have allowed more time for excess deaths to occur. The importance of latency in detecting excess mortality as an outcome was highlighted in a recent study from The Netherlands, which identified increased mortality in an incident cohort of RA patients (symptom duration <1 year at baseline) only after 10 years of

Table 3. Poisson regression model by cohort after 7 years of followup*

	MRR (95% CI), unadjusted				MRR (95% CI), adjusted†			
	EIA	2010 RA criteria positive	2010 RA criteria negative	RF/ACPA positive	EIA	2010 RA criteria positive	2010 RA criteria negative	RF/ACPA positive
All cause								
Cohort 1	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Cohort 2	0.97 (0.77–1.23)	1.10 (0.81–1.49)	0.89 (0.55–1.46)	0.81 (0.55–1.20)	0.97 (0.77–1.24)	1.13 (0.84–1.52)	0.89 (0.54–1.49)	0.82 (0.56–1.19)
Cohort 3	0.88 (0.67–1.15)	1.01 (0.71–1.45)	0.84 (0.49–1.46)	0.90 (0.58–1.38)	0.89 (0.68–1.17)	1.00 (0.70–1.43)	0.84 (0.47–1.51)	0.89 (0.58–1.35)
Cardiovascular								
Cohort 1	Ref.	Ref.	–‡	–‡	Ref.	Ref.	–‡	–‡
Cohort 2	0.92 (0.62–1.37)	1.02 (0.61–1.69)	–‡	–‡	0.94 (0.63–1.39)	1.07 (0.65–1.76)	–‡	–‡
Cohort 3	0.88 (0.56–1.39)	1.07 (0.58–1.99)	–‡	–‡	0.93 (0.59–1.46)	1.08 (0.58–1.98)	–‡	–‡

* MRR = mortality rate ratio; 95% CI = 95% confidence interval; EIA = early inflammatory arthritis; RA = rheumatoid arthritis; RF = rheumatoid factor; ACPA = anti–citrullinated protein antibody.
† Adjusted for age at symptom onset and sex and symptom duration at baseline.
‡ Too few events to calculate the MRR.

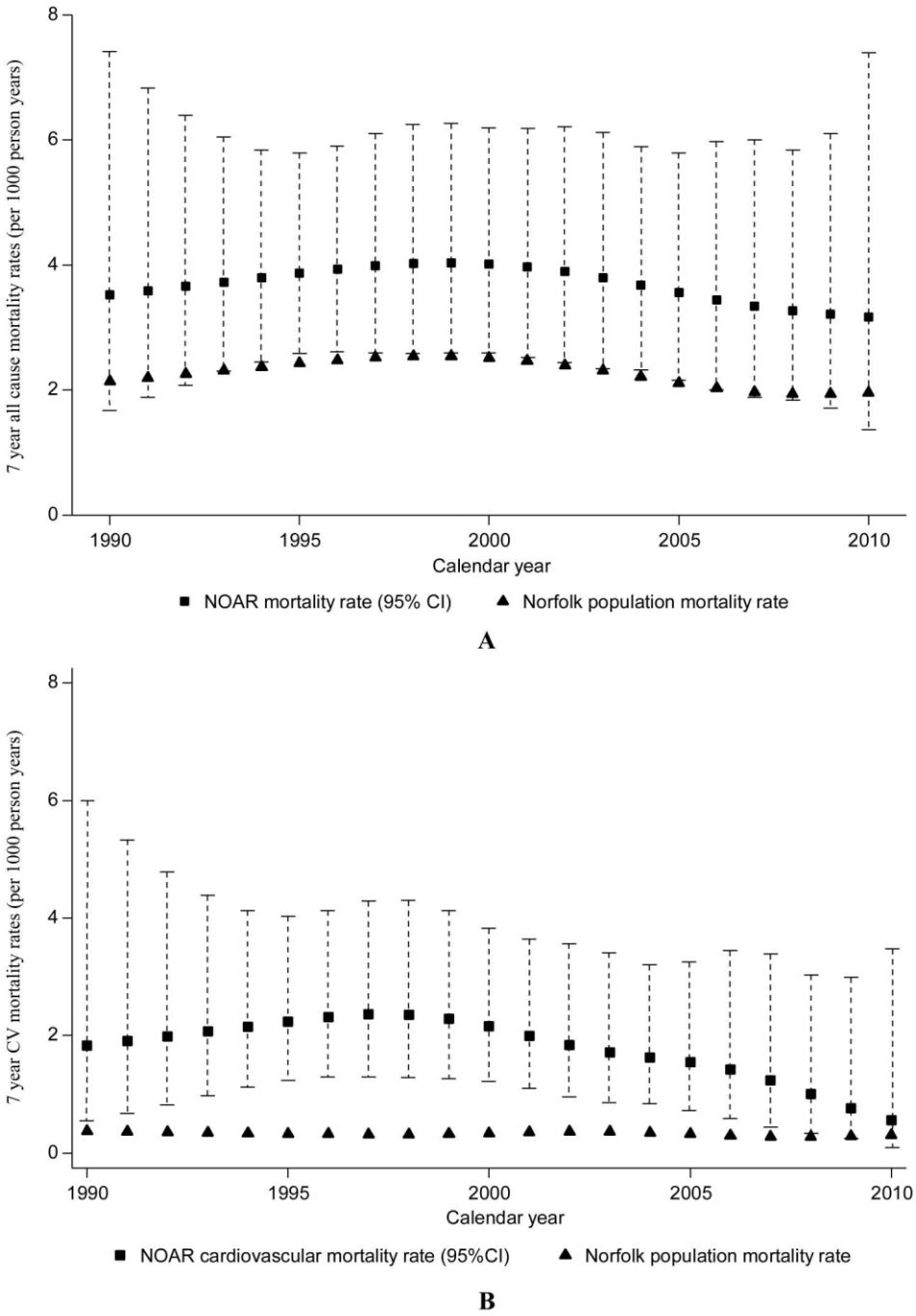


Figure 1. Observed and expected mortality rates by calendar year for all-cause (**A**) and cardiovascular (CV)–specific (**B**) mortality. Observed rates were modeled using Poisson regression with natural splines and expected rates were calculated from mortality rates for Norfolk, age and sex standardized to the study population. The y-axes show rates per 1,000 person-years and 95% confidence intervals (95% CIs). NOAR = Norfolk Arthritis Register.

followup (3). We also found no decrease in mortality rates over time after accounting for trends in the background population. These findings are consistent with a recent meta-analysis by Dadoun et al (2), who collated 8 studies reporting SMRs of patients with early RA (<2 years' duration) published in 1955–1995. They found that mortality rates in RA patients remain elevated compared to the general population, and had not altered significantly over time. It remains to be seen whether even more aggressive remission-targeted therapy will alter this, whether there will be an impact on mortality in the longer term, or whether, for example, antibody status is an unmodifiable risk factor for decreased survival in patients with IA.

In keeping with previous results from the NOAR (11,12) we found that, in the subgroup of patients who were ACPA or RF positive, there were 40% more deaths than expected. This proportion was higher than in patients who met the RA classification criteria and suggests antibody status

plays an important role in the increased mortality seen in RA. RF is an established risk factor for increased mortality in RA (1) and even has been identified as a risk factor in subjects without joint symptoms (16). Since ACPA testing has only been routinely available in the past 5–10 years, the literature examining the relationship between ACPA and mortality in RA is limited. However, a similar association with RF appears to exist (17).

There are limitations to this study. Although SMRs are a widely used measure of mortality risk, comparisons between SMRs measured in different cohorts and time periods must be made with caution. This is because the expected number of deaths is dependent on the length of followup, the age and sex structure of the disease cohort, and the mortality rates in the general population. Although we kept the period of followup constant between the cohorts, the age at onset of EIA increased during the period of the study, and so the expected number of deaths will have risen. We used MRRs to make comparisons between the cohorts and modeled the rates using Poisson regression, adjusted for age at onset and sex and symptom duration at presentation, to allow for these differences.

In conclusion, we have shown that mortality in EIA remains elevated compared to the general population, and mortality rates have not changed significantly over the past 20 years. We have demonstrated, for the first time, increased SMRs in patients satisfying the 2010 ACR/EULAR classification criteria for RA, and further demonstrated the importance of autoantibody status in the excess mortality seen in patients with IA.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Verstappen had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Humphreys, Warner, Marshall, Symmons, Verstappen.

Acquisition of data. Chipping, Marshall.

Analysis and interpretation of data. Humphreys, Warner, Lunt, Symmons, Verstappen.

REFERENCES

1. Sokka T, Abelson B, Pincus T. Mortality in rheumatoid arthritis: 2008 update. Clin Exp Rheumatol 2008;26 Suppl: S35–61.
2. Dadoun S, Zeboulon-Ktorza N, Combescure C, Elhai M, Rozenberg S, Gossec L, et al. Mortality in rheumatoid arthritis over the last fifty years: systematic review and meta-analysis. Joint Bone Spine 2013;80:29–33.
3. Radovits BJ, Fransen J, Al Shamma S, Eijsbouts AM, van Riel PL, Laan RF. Excess mortality emerges after 10 years in an inception cohort of early rheumatoid arthritis. Arthritis Care Res (Hoboken) 2010;62:362–70.
4. Smolina K, Wright FL, Rayner M, Goldacre MJ. Determinants of the decline in mortality from acute myocardial infarction in England between 2002 and 2010: linked national database study. BMJ 2012;344:d8059.
5. Sugiyama D, Nishimura K, Tamaki K, Tsuji G, Nakazawa T, Morinobu A, et al. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. Ann Rheum Dis 2010;69:70–81.
6. Health & Social Care Information Centre. Statistics on smoking: England, 2011. 2011. URL: http://www.hscic.gov.uk/.
7. Saunders SA, Capell HA, Stirling A, Vallance R, Kincaid W, McMahon AD, et al. Triple therapy in early active rheumatoid arthritis: a randomized, single-blind, controlled trial comparing step-up and parallel treatment strategies. Arthritis Rheum 2008;58:1310–7.
8. Gonzalez A, Maradit Kremers H, Crowson CS, Nicola PJ, Davis JM III, Therneau TM, et al. The widening mortality gap between rheumatoid arthritis patients and the general population. Arthritis Rheum 2007;56:3583–7.
9. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
10. Harrison BJ, Symmons DP, Barrett EM, Silman AJ. The performance of the 1987 ARA classification criteria for rheumatoid arthritis in a population based cohort of patients with early inflammatory polyarthritis. J Rheumatol 1998;25:2324–30.
11. Goodson NJ, Wiles NJ, Lunt M, Barrett EM, Silman AJ, Symmons DP. Mortality in early inflammatory polyarthritis: cardiovascular mortality is increased in seropositive patients. Arthritis Rheum 2002;46:2010–9.
12. Naz SM, Farragher TM, Bunn DK, Symmons DP, Bruce IN. The influence of age at symptom onset and length of followup on mortality in patients with recent-onset inflammatory polyarthritis. Arthritis Rheum 2008;58:985–9.
13. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.
14. Symmons DP, Barrett EM, Bankhead CR, Scott DG, Silman AJ. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. Br J Rheumatol 1994;33:735–9.
15. World Health Organization. International Statistical Classification of Diseases and Related Health Problems, tenth revision (ICD-10). 4th ed. Geneva: WHO; 2010.
16. Tomasson G, Aspelund T, Jonsson T, Valdimarsson H, Felson DT, Gudnason V. Effect of rheumatoid factor on mortality and coronary heart disease. Ann Rheum Dis 2010;69:1649–54.
17. Farragher TM, Goodson NJ, Naseem H, Silman AJ, Thomson W, Symmons D, et al. Association of the HLA–DRB1 gene with premature death, particularly from cardiovascular disease, in patients with rheumatoid arthritis and inflammatory polyarthritis. Arthritis Rheum 2008;58:359–69.

5.3.1 Supplementary results

There were no missing data in the analysis in paper 5.3 and no supplementary data files.

Chapter 6: Results

Longitudinal outcomes

This chapter comprises 2 manuscripts. The first analyses whether patients with early inflammatory arthritis who satisfy the 2010 ACR/EULAR classification criteria have more disability, disease activity and radiographic damage over up to 20 years follow up than those patients who do not satisfy the criteria. The second manuscript analyses whether a novel family of autoantibodies, the anti-carbamylated protein antibodies, are associated with greater disability and disease activity, also over up to 20 years follow up.

6 Results: Longitudinal outcomes

6.1 The 2010 ACR/EULAR classification criteria predict disability, disease activity and radiographic damage over long term follow up of patients with early inflammatory arthritis: results from the Norfolk Arthritis Register

These results have not been submitted for publication but are presented in manuscript format for consistency.

The 2010 ACR/EULAR classification criteria predict disability, disease activity and radiographic damage over long term follow up of patients with early inflammatory arthritis: results from the Norfolk Arthritis Register

Jennifer H Humphreys[1], Jackie Chipping[2], Deborah PM Symmons[1,3], Suzanne MM Verstappen[1]

[1]Arthritis Research UK Centre for Epidemiology, University Of Manchester, Manchester
[2]Norfolk and Norwich University Hospital, Norwich, United Kingdom
[3] NIHR Manchester Musculoskeletal Biomedical Research Unit, Manchester Academic Health Science Centre, Manchester, UK

Key word: Early inflammatory arthritis, rheumatoid arthritis, classification criteria, disability, disease severity, erosions

Word count (including abstract):3050

Correspondence to:

Dr Suzanne M.M. Verstappen, Arthritis Research UK Centre for Epidemiology, Manchester Academic Health Science Centre, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, UK

Email: suzanne.verstappen@manchester.ac.uk

Tel: +44 (0)161 275 5663

Abstract

Aims: To investigate the association between the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria for rheumatoid arthritis (RA) and disability, disease activity and radiographic damage over up to 20 years follow up.

Methods: Adults with ≥2 swollen joints for ≥ 4 weeks, with symptom duration of <2 years, were recruited to the Norfolk Arthritis Register (NOAR) from 1990-2009. At baseline, a research nurse took details of patient demographics, smoking status and performed a tender and swollen joint count. A blood sample was taken and rheumatoid factor (RF), anti-citrullinated protein antibody (ACPA) and C-reactive protein (CRP) were measured. Patients completed the Health Assessment Questionnaire (HAQ) to assess baseline disability. Classification criteria were applied and disease activity measured using the DAS28 score. HAQ and DAS28 scores were repeated at regular intervals over follow up. Radiographs were taken of hands and feet and scored using the Larsen method at baseline, 1st or 2nd anniversary, and repeated in some patients at 5th anniversary. Generalised estimating equations were used to analyse the association between satisfying the 2010 ACR/EULAR classification criteria and longitudinal HAQ, DAS28 and erosive disease. Zero inflated negative binomial regression was used to analyse the association between satisfying the 2010 criteria and the Larsen score. Subgroup analyses were performed in patients recruited before and after the year 2000. The proportion of 2010 criteria negative patients who went on to require DMARDs and had HAQ scores greater than 1 were described.

Results: 3054 patients were included, 1723 (63%) satisfied the 2010 criteria at baseline. Patients who satisfied the criteria had greater burden of disability throughout follow up, β (95% confidence interval (CI)) 0.38 (0.33, 0.43). Similar results were seen for DAS28: β (95% CI) 1.63 (1.54, 1.73), the presence of erosions: odds ratio (95%CI) 1.99 (1.65, 2.41) and the Larsen score: β (95% CI) 0.33 (0.20, 0.47). Results were comparable in the pre and post 2000 subgroups, although the increase in likelihood of erosions and burden of radiographic damage was larger in pre-2000 subgroup. The proportions of patients with 2010 criteria negative IP who required DMARDS and had moderate to severe disability by 5 years follow up were 24 and 25% respectively.

Conclusions: The 2010 criteria predict long term disability, disease activity and radiographic damage. The similar associations seen in patients recruited before and after the availability of improved disease management, except for radiographic damage, suggest the criteria are able to identify those with worst future outcomes, despite modern therapeutic interventions. However patients who do not fulfil the criteria may still have poor outcomes. These findings provide important construct validity for the criteria.

Background:

The ACR/EULAR classification criteria for rheumatoid arthritis (RA) (1) have now been available to clinicians and researchers for nearly 5 years. They have been shown to classify more patients early in their disease course (2) and to have improved sensitivity over the 1987 criteria in identifying early those patients with inflammatory arthritis who go on to require intervention with disease-modifying therapy (3), and persistent arthritis (4). In addition, patients presenting with inflammatory arthritis who meet the classification criteria are more likely to die earlier than those who do not (5). However, there are few data currently describing the impact of satisfying the criteria on long term disability, disease activity and radiographic damage. Long term outcomes can be difficult to interpret in the context of modern intensive treatment strategies and biologic therapies for RA. It would be useful to study the impact of meeting the criteria in historic datasets from an era when disease modifying anti-rheumatic drugs (DMARDs) were used less frequently and biologic drugs were not available. The first aim of this study, therefore, was to investigate the burden of disability, disease activity and radiographic damage over up to 20 years of follow up in a cohort of patients with early inflammatory arthritis (EIA) who satisfy the 2010 criteria compared to those who do not. The second aim was to examine these associations in the subgroups of these patients recruited before and after the year 2000. The final aim was to describe outcomes in patients who did not satisfy the 2010 criteria.

Methods:

Patients and setting: Patients were recruited to the Norfolk Arthritis Register (NOAR) from 1990-2009 if they presented to primary or secondary care with at least 2 swollen joints for at least 4 weeks. For the purpose of this analysis patients were only included if they presented within 2 years of the onset of symptoms. At baseline, a research nurse took details of patient demographics, comorbidities, smoking status and performed a 51 tender and swollen joint count. Patients completed the Health Assessment Questionnaire (HAQ) (6) and blood was taken and serum stored frozen. This was subsequently tested for rheumatoid factor (RF) (latex test), anti-citrullinated protein antibodies (Axis-Shield Diastat Anti-CCP kit) and C-reactive protein (CRP). The 3 item disease activity score (DAS28-CRP) was calculated (7). Patients in NOAR are followed up annually for the first 3 years, then at 5, 7, 10, 15 and 20 years from baseline. Patients repeat the HAQ and the nurse assessment at each follow up and blood samples are taken every 5 years allowing completion of DAS28-

CRP. Radiographs of hands and feet are taken in all patients who met the 1987 ACR classification criteria at 1st or 2nd follow up if they had symptom onset before the year 2000 and at baseline and 1st follow up if their symptom onset occurred after 2000. Additional radiographs were obtained at other follow up assessments where a patient could satisfy the 1987 RA classification criteria if radiographic information were available. Larsen scoring by two independent readers was used to measure radiographic damage.

Statistical analysis: Cross-sectional differences at baseline between patients who did and did not satisfy the 2010 criteria were examined using univariate linear and logistic regression for continuous and categorical variables respectively. Generalised estimating equations (GEE) were used to investigate the association between satisfying the 2010 criteria and the outcomes of interest. For HAQ and DAS28 this was based on a standard linear regression; for the presence or absence of erosions it was based on logistic regression. As Larsen scores usually follow a highly non-normal distribution with a large proportion of patients scoring zero, zero-inflated negative binomial regression was used, which also allows for the irregular way in which radiographic data were obtained over time. Multivariate models adjusted for age at symptom onset, gender, disease duration, baseline smoking status and year recruited to the register. The analyses were repeated in 2 subgroups dividing the patients into those recruited prior to the year 2000 and those recruited after that time. Sensitivity analyses were conducted using random effects models to confirm the associations.

To investigate whether patients who did not fulfil the 2010 criteria ever had poor outcomes, two descriptive analyses were performed. Patients were only included in these analyses if they were DMARD naïve at baseline assessment. This is because being on treatment could reduce the likelihood of satisfying the 2010 criteria. At each follow up, the proportion of patients with EIA that had subsequently been prescribed DMARDs and the proportion with HAQ score greater than 1 were calculated, using the total number who did not satisfy the 2010 criteria as the denominator.

Results:

A total of 3054 patients were included in this analysis, of whom 1723 (56%) satisfied the 2010 criteria. Cohort and subgroup characteristics along with the prevalence of missing data are shown in table one. Patients who satisfied the 2010 criteria were slightly more

likely to be female, odds ratio (OR) (95% confidence interval (CI)) 1.52 (1.41, 1.63). They were slightly older (β coefficient (95% CI) 0.98 (0.45, 1.5)) and had slightly longer symptom duration at presentation (β coefficient (95% CI) 3.5 (2.72, 4.27)). RF and ACPA were present in 780 (49%) and 539 (42%) of patients who met the criteria respectively, but were rarely found in patients who did not meet the criteria. In addition patients who satisfied the criteria at baseline had higher levels of CRP, DAS28 and HAQ. There was no difference in the proportions of current, ex- or never smokers between the two groups. Median follow up time (interquartile range, IQR) from baseline to most recent assessment was 5.5 (2.3-9.9) years in the total cohort; in the 2010 criteria positive and negative groups it was 6.8 (3.1-9.9) years and 5.1 (1.8-9.8) years respectively.

In the GEE analyses, patients who satisfied the 2010 criteria had more disability and higher levels of disease activity throughout follow up, respective adjusted β coefficients (95% CI) 0.38 (0.33, 0.43) and 1.63 (1.54, 1.73), see table 2. In the subgroup analyses, patients who met the classification criteria and were recruited to NOAR before the year 2000 had slightly higher disease activity scores on average over follow up than those recruited from the year 2000 onwards, adjusted β coefficients (95% CIs) 1.71 (1.60, 1.83) and 1.57 (1.41, 1.73) respectively (table 2). In terms of disability, patients who satisfied the 2010 criteria and were recruited earlier had slightly lower levels of disability compared to those who were recruited later, adjusted β coefficients (95% CIs) 0.38 (0.32, 0.44) and 0.42 (0.33, 0.50) respectively. Similar results were seen in the random effects models (supplementary table 1).

Table 1. Cohort characteristics

	Total cohort n=3054	2010 criteria positive n=1723	2010 criteria negative n=1331	pre2000 n=1913	post2000 n=1141	Missing at baseline
Female	1971 (65)	1167 (68)	804 (60)	1241 (65)	730 (64)	0
Age onset	56 (44-68)	57 (46-68)	54 (40-67)	55 (42-67)	57 (46-69)	0
Symptom duration (weeks)	26 (14-47)	27 (16-48)	25 (13-45)	25 (13-46)	28 (17-48)	0
ACPA positive	609 (27)	539 (42)	70 (0.1)	369 (25)	240 (32)	822 (27)
RF positive	912 (33)	780 (49)	132 (0.1)	478 (28)	434 (42)	331 (11)
HAQ score	0.75 (0.25-1.5)	1.125 (0.5-1.75)	0.375 (0-1)	0.75 (0.25-1.375)	0.875 (0.375-1.5)	36 (1)
DAS28	3.71 (2.78-4.76)	4.47 (3.62-5.36)	2.72 (2.09-3.37)	3.74 (2.75-4.85)	3.67 (2.80-4.62)	561 (18)
CRP	9 (2-20)	11 (4-24)	5 (0-13)	7 (0-17)	11 (5-22)	561 (18)
Smoking status:						24 (0.01)
current	753 (25)	449 (26)	303 (27)	494 (26)	259 (23)	
previous	1192 (39)	691 (40)	499 (44)	744 (39)	448 (39)	
never	1085 (36)	574 (33)	332 (29)	664 (35)	421 (37)	
median follow up time (years)	5.5 (2.3-9.9)	6.8 (3.1-9.9)	5.1 (1.8-9.8)	5.2 (1.5-14.2)	6.8 (4.9-9.8)	-

ACPA, anti-citrullinated protein antibodies; RF, rheumatoid factor; HAQ, health assessment questionnaire; DAS28, disease activity score; CRP, C-reactive protein

Table 2. Association between classification criteria and HAQ and DAS28 scores

	Total cohort *β (95% CI)*	**Pre2000** *β (95% CI)*	**Post2000** *β (95% CI)*
HAQ			
Univariate	0.46 (0.41,0.51)	0.48 (0.43,0.54)	0.41 (0.33,0.49)
Multivariate*	0.38 (0.33,0.43)	0.38 (0.32,0.44)	0.42 (0.33,0.50)
DAS28			
Univariate	1.27 (1.19,1.35)	1.29 (1.19,1.39)	1.25 (1.12,1.38)
Multivariate*	1.63 (1.54,1.73)	1.71 (1.60,1.83)	1.57 (1.41,1.73)

* adjusted for age at symptom onset, gender, disease duration, baseline smoking status and year recruited to the register

The timing of the radiographic examinations taken during follow up in the subgroups of patients recruited before and after year 2000 is shown in table 3. A total of 1990 patients had one or more sets of radiographs over the follow up period, of whom 834 patients had at least 2 sets of radiographs.

Table 3. Timing of radiographic examinations

Follow up year of radiographs	**Number of radiographs** *n* (% cohort/subgroup)*		
	Total cohort *n=3054*	**Pre- 2000** *n=1913*	**Post 2000** *n=1141*
0	561(18)	212(11)	349(31)
1	1171(38)	787(41)	348(30)
2	375(12)	297(16)	78(0.07)
3	20(0.01)	2(0.001)	18(0.02)
5	826(27)	731(38)	95(0.08)
Total	2953	2029	888

*834 patients had radiographs performed at multiple follow up assessments

The 2010 criteria were associated with greater overall radiographic damage as measured by the Larsen score in the zero inflated negative binomial regression model, β coefficient (95% CIs) 0.33 (0.20, 0.47). The criteria were also associated with higher likelihood of developing erosions, odds ratio (95%CI) 1.99 (1.65, 2.41), see table 4. These associations were

maintained in both the early and later cohorts, but were stronger in patients recruited before the year 2000.

Table 4. Association between classification criteria and radiographic damage

	Total cohort *β (95% CI)*	**Pre2000** *β (95% CI)*	**Post2000** *β (95% CI)*
Larsen Score			
Univariate	0.40 (0.27,0.53)	0.47 (0.31,0.63)	0.16 (-0.07,0.39)
Multivariate*	0.33 (0.20,0.47)	0.47 (0.31,0.63)	0.23 (0.02,0.44)
Erosions (yes/no)	*OR(95% CI)*	*OR(95% CI)*	*OR(95% CI)*
Univariate	2.00 (1.67,2.39)	2.22 (1.77,2.79)	1.72 (1.29,2.30)
Multivariate*	1.99 (1.65,2.41)	2.39 (1.89,3.05)	1.91 (1.38,2.63)

*adjusted for age at symptom onset, gender, disease duration, baseline smoking status and year recruited to the register

Table 5 shows the number of 2010 criteria negative patients at each follow up who had received DMARD therapy and had HAQ scores greater than 1. A subgroup of patients who did not fulfil the 2010 criteria did experience these poor outcomes. For example, a total of 629 patients who did not satisfy the 2010 criteria had 5 year follow up data, of these 187 (30%) had been treated with DMARDs and 193 (31%) had HAQ scores over 1.

Table 5. Poor outcomes in 2010 criteria negative patients by year of follow up*

Year of follow up	**No. of 2010 criteria negative patients**[$]	**No. with HAQ score >1** *n(%)*[$]	**No. treated with DMARDs** *n(%)*[$]
0	1015	233 (23)	0*
1	877	219 (25)	189 (22)
2	662	167 (25)	181 (27)
3	661	177 (27)	196 (30)
5	629	193 (31)	187 (30)
7	414	127 (31)	138 (33)
10	271	93 (34)	82 (30)
15	166	56 (34)	46 (28)
20	99	40 (40)	24 (24)

*Patients were only included in this table if they had not yet received DMARDs at the time of the baseline assessment
[$] % of the total number of baseline 2010 criteria negative patients (who were also DMARD naïve at baseline) at that follow up

Discussion:

In this study we have demonstrated that patients with EIA who fulfil the 2010 criteria have more long term disability, higher levels of disease activity and greater radiographic damage compared to those who do not. These findings are important for a number of reasons. Firstly, disability is a key outcome for patients themselves, as it describes the impact of RA on their lives. Secondly, although the 2010 criteria were not recommended to be used as a diagnostic tool or intended to guide therapeutic decisions (1), whether or not a patient satisfies classification criteria can be highly influential for the treating physician. Finally, classification criteria are correctly used as entry criteria for clinical trials and other research studies, and their association with long term outcomes is useful knowledge in that context. Nevertheless, in this cohort we also identified a proportion of patients with EIA who did not satisfy the 2010 criteria but went on to require intervention with DMARDs and experience moderate to severe levels of disability. Therefore, it is clear that patients who do not satisfy the 2010 criteria do not have a universally good prognosis.

In addition we were able to describe differences in outcomes dependent on when patients were recruited to our study. In recent times, it has been more challenging to investigate long term outcomes related to cumulative disease activity in RA. This is because highly effective biologic therapies and aggressive treat-to-target management strategies aim to minimise disease activity (8-10). Different approaches can be employed to try to deal with this problem, for example complex statistical modelling such as propensity scoring and marginal structural modelling, however there always remains potential for residual confounding (11;12). Alternatively patients could be followed for longer to allow them to accumulate sufficient damage, although it is difficult to know how much longer they would need to be observed. An historic cohort such as NOAR offers an opportunity for another approach. It allows us to study the natural progression of these outcomes in patients recruited before newer treatments and strategies were available or in common use, and as a result the associations we have demonstrated in the early cohort are not influenced by them.

When the patients were divided by era, the most striking differences were seen in the radiographic data. The association between satisfying the criteria and the burden of radiographic damage as measured by the Larsen score was much stronger in the group recruited earlier. This may in part be due the different ways in which radiographic data were routinely collected. Similarly, in terms of erosions, patients who satisfied the criteria

were more likely to develop erosions in both subgroups, but the odds were greater in patients recruited before the year 2000. This may suggest that modern treatment strategies targeting remission are resulting in fewer erosions. Interestingly, however, the findings for disability and disease activity showed no marked differences between the two eras. In fact, the association between satisfying the criteria and disability during follow up was slightly greater in patients recruited more recently. This may be due to changes in baseline disability, as it has been shown previously in patients recruited to NOAR that baseline disability has increased over time (13), and this therefore may reflect disability not related to their arthritis. It also suggests that while clinicians need to continue to target low disease activity and remission, they should also consider other factors recognised to contributing to patient disability, such as obesity (14) and comorbidity (15), in order to achieve the best outcomes for our patients. It is important to note, however, that the differences in outcomes between the two eras are reported descriptively and the statistical comparisons are made within each group between patients with EIA who do and do not satisfy the criteria.

Since their publication in 2010, many studies have investigated to short and medium term outcomes of patients who fulfil the classification criteria (3), frequently focussing on whether these patients are initiated on DMARDs (16-18) or receive a clinical diagnosis of RA (17;18). These studies provide some measure of external validity but these results only relate to the initial phase of diagnosis and management. Patients may be more concerned about how RA will impact on their lives many years in the future. However relatively few groups have been able to investigate outcomes over extended follow up, which would provide a more robust idea of the construct validity of the criteria. In the SCQM study these outcomes were addressed in the medium term (19); over a median of 3.6 years follow up they found that patients with EIA who fulfilled the 2010 criteria at baseline had higher levels of disease activity and more radiographic progression than those who did not, although therapeutic strategies did not differ significantly between the two groups (19). A study from Finland (21) investigated the ability of the 2010 criteria to differentiate RA from other inflammatory arthritides, using diagnosis 10 years later as the gold standard, and found that only 4% of the non-RA patients were misclassified as RA by the criteria. Our study remains unique, however, in the length of follow up data available for the patients recruited into the early cohort.

Other groups have made comparisons between the long term outcomes of patients meeting the 1987 criteria and those who meet the 2010 criteria. In a large study of over

1500 patients with up to 10 years follow up from the Leiden Early Arthritis Clinic, patients who fulfilled the 2010 criteria had less severe radiographic damage and were more likely to achieve DMARD free remission than those who fulfilled the 1987 criteria (20). A multicentre study from France followed 164 patients with EIA for 10 years, and found that the 2010 criteria had similar sensitivity but higher specificity than the 1987 criteria to predict a physician diagnosis of RA by the end of follow up.

There are some limitations to our study. There were some missing data in the cohort. The regression models used in this study allow patients to have gaps in follow up, and sensitivity analyses conducted using random effects models, which allow for missing at random data, produced similar results. However there were also some missing baseline data, particularly regarding serum samples. The 2010 criteria recommend that where a data item is missing, it should be treated as zero (22) and we have followed that recommendation. However it is possible that some patients could have been classified as having RA if more information were available. The effect of this on our analysis would be to produce a more conservative estimate of the association between satisfying the classification criteria and the outcomes of interest. The 2010 criteria also allow patients to be classified as RA if they have evidence of typical RA erosions on radiographs taken at the time of classification, defined as an erosion in at least three separate joints at any of the following sites: the proximal interphalangeal, the metacarpophalangeal, the wrist and the metatarsophalangeal joints (23). In NOAR baseline radiographs were not taken routinely in all patients and we have therefore not applied this criterion. Again this is likely to have led to an underestimation of the association between the 2010 criteria and outcomes. In addition, it should be highlighted that the poor outcomes in patients who did not satisfy the 2010 criteria, and the differences between the two eras were simple descriptive analyses, and no statistical inferences were made.

In conclusion, this study to demonstrated significant associations between satisfying the 2010 classification criteria for RA and disability and disease activity over up to 20 years, and that these associations appear to be independent of modern treatments. Nevertheless a proportion of patients who do not satisfy the criteria continue to experience poor outcomes. These results will be important for both patients presenting with EIA and their healthcare workers to inform their long term disease management and care.

Reference List

(1) Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010 Sep;69(9):1580-8.

(2) Humphreys JH, Verstappen SM, Hyrich KL, Chipping JR, Marshall T, Symmons DP. 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 2012 Sep 3.

(3) Humphreys JH, Symmons DP. Postpublication validation of the 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: where do we stand? Curr Opin Rheumatol 2013 Mar;25(2):157-63.

(4) van der Linden MP, Knevel R, Huizinga TW, van der Helm-van Mil AH. Classification of rheumatoid arthritis: comparison of the 1987 American College of Rheumatology criteria and the 2010 American College of Rheumatology/European League Against Rheumatism criteria. Arthritis Rheum 2011 Jan;63(1):37-42.

(5) Humphreys JH, Verstappen SMM, Hyrich KL, Chipping J, Symmons DP. 2010 ACR/EULAR classification criteria for rheumatoid arthritis predict increased mortality in patients with early arthritis: results from the Norfolk Arthritis Register. Rheumatology (Oxford, England) 2013;52(6):1141-2.

(6) Wolfe F, Kleinheksel SM, Cathey MA, Hawley DJ, Spitz PW, Fries JF. The clinical value of the Stanford Health Assessment Questionnaire Functional Disability Index in patients with rheumatoid arthritis. J Rheumatol 1988 Oct;15(10):1480-8.

(7) Department of Rheumatology UMCNN. DAS28. http://www das-score nl/ 2004Available from: URL: http://www.das-score.nl/

(8) Boers M, Verhoeven AC, Markusse HM, van de Laar MA, Westhovens R, van Denderen JC, et al. Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. Lancet 1997 Aug 2;350(9074):309-18.

(9) Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, Van ZD, Kerstens PJ, Hazes JM, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. Arthritis Rheum 2005 Nov;52(11):3381-90.

(10) Nam JL, Ramiro S, Gaujoux-Viala C, Takase K, Leon-Garcia M, Emery P, et al. Efficacy of biological disease-modifying antirheumatic drugs: a systematic literature review informing the 2013 update of the EULAR recommendations for the management of rheumatoid arthritis. Ann Rheum Dis 2014 Mar;73(3):516-28.

(11) Mercer LK, Lunt M, Low AL, Dixon WG, Watson KD, Symmons DP, et al. Risk of solid cancer in patients exposed to anti-tumour necrosis factor therapy: results from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis. Ann Rheum Dis 2014 Mar 31.

(12) Yang S, Eaton CB, McAlindon TE, Lapane KL. Effects of glucosamine and chondroitin on treating knee osteoarthritis: An analysis with marginal structural models. Arthritis Rheumatol 2014 Nov 4.

(13) Diffin JG, Lunt M, Marshall T, Chipping JR, Symmons DP, Verstappen SM. Has the severity of rheumatoid arthritis at presentation diminished over time? J Rheumatol 2014 Aug;41(8):1590-9.

(14) Humphreys JH, Verstappen SM, Mirjafari H, Bunn D, Lunt M, Bruce IN, et al. Association of morbid obesity with disability in early inflammatory polyarthritis: results from the Norfolk Arthritis Register. Arthritis Care Res (Hoboken) 2013 Jan;65(1):122-6.

(15) Norton S, Koduri G, Nikiphorou E, Dixey J, Williams P, Young A. A study of baseline prevalence and cumulative incidence of comorbidity and extra-articular manifestations in RA and their impact on outcome. Rheumatology (Oxford) 2013 Jan;52(1):99-110.

(16) Cader MZ, Filer A, Hazlehurst J, de PP, Buckley CD, Raza K. Performance of the 2010 ACR/EULAR criteria for rheumatoid arthritis: comparison with 1987 ACR criteria in a very early synovitis cohort. Ann Rheum Dis 2011 Jun;70(6):949-55.

(17) Britsemmer K, Ursum J, Gerritsen M, van Tuyl LH, Van SD. Validation of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: slight improvement over the 1987 ACR criteria. Ann Rheum Dis 2011 Aug;70(8):1468-70.

(18) Biliavska I, Stamm TA, Martinez-Avila J, Huizinga TW, Landewe RB, Steiner G, et al. Application of the 2010 ACR/EULAR classification criteria in patients with very early inflammatory arthritis: analysis of sensitivity, specificity and predictive values in the SAVE study cohort. Ann Rheum Dis 2012 Sep 15.

(19) Mueller RB, Schiff M, Kaegi T, Finckh A, Haile SR, Schulze-Koops H, et al. The new 2010 ACR/EULAR criteria as predictor of clinical and radiographic response in patients with early arthritis. Clin Rheumatol 2015 Jan;34(1):51-9.

(20) Burgers LE, van Nies JA, Ho LY, de Rooy DP, Huizinga TW, van der Helm-van Mil AH. Long-term outcome of rheumatoid arthritis defined according to the 2010-classification criteria. Ann Rheum Dis 2014 Feb;73(2):428-32.

(21) Kaarela K, Kauppi JE, Kauppi MJ. The 2010 ACR/EULAR classification criteria for rheumatoid arthritis in the Heinola inception cohort--diagnoses confirmed by long-term follow-up. Clin Rheumatol 2012 Mar;31(3):547-51.

(22) Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, III, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010 Sep;62(9):2569-81.

(23) van der Heijde D, van der Helm-van Mil AH, Aletaha D, Bingham CO, Burmester GR, Dougados M, et al. EULAR definition of erosive disease in light of the 2010 ACR/EULAR rheumatoid arthritis classification criteria. Ann Rheum Dis 2013 Apr;72(4):479-81.

6.1.1 Supplementary results

6.1.2 Differences between patients with missing and non-missing baseline data in results section 6.1

Although there were missing data on all variables which included required blood samples, the largest proportion of missing was in ACPA. The data were therefore inspected for differences between patients with and without baseline ACPA status; these are shown in table 6.1.1. The characteristics of these two groups were similar and univariate analyses of differences between the two groups are reported in the main manuscript in section 6.1.

Table 6.1.1 Baseline characteristics of patients with and without ACPA status

	Missing baseline ACPA *n=822*	Non-missing *n=2232*
Female	530(64)	1441(65)
Age onset	56(43-70)	55(44-67)
Symptom duration *(weeks)*	25(14-42)	26(15-48)
HAQ score	0.875(0.25-1.5)	0.75(0.25-1.375)
DAS28	3.69(2.82-4.73)	3.71(2.75-4.76)
CRP	11(3-25)	8(1-19)
Smoking status:		
current	209(26)	544(24)
previous	314(39)	878(39)
never	284(35)	801(36)
Satisfy 2010 RA criteria	437(53)	1286(58)
Follow up time *(years)*	6.9(2.7-10.0)	5.4(2.3-9.9)

Categorical variables are reported as n(%). Continuous variables are reported as median (interquartile range)

HAQ: Health Assessment Questionnaire, DAS28: disease activity score with 28 joints, CRP: C-reactive protein

6.1.3 Sensitivity analyses

Table 6.1.2 show the results of the random effects models, constructed as a sensitivity analysis for the GEE models testing the association between patients with IP who satisfy the 2010 criteria at baseline and HAQ and DAS28 scores over time. These show very similar estimates to those seen in the GEE models in paper 6.1, thus adding to the robustness of the results.

Table 6.1.2 Random effects model of the association between 2010 classification criteria and HAQ and DAS28

	Total cohort *β (95% CI)*	Pre2000 *β (95% CI)*	Post2000 *β (95% CI)*
HAQ			
Univariate	0.46(0.41-0.51)	0.48(0.42-0.55)	0.41(0.33-0.49)
Multivariate*	0.38(0.34-0.43)	0.39(0.33-0.45)	0.42(0.34-0.50)
DAS28			
Univariate	1.27(1.19-1.35)	1.28(1.18-1.38)	1.25(1.12-1.37)
Multivariate*	1.62(1.53-1.72)	1.69(1.57-1.81)	1.56(1.41-1.72)

*adjusted for age at symptom onset, gender, disease duration, baseline smoking status and year recruited to the register

HAQ: Health Assessment Questionnaire, DAS28: disease activity score with 28 joints,

6.2 Anti-carbamylated protein antibodies are associated with long term disability and increased disease activity in patients with early inflammatory arthritis: results from the Norfolk Arthritis Register (NOAR)

Currently in press, accepted by the Annals of the Rheumatic Diseases.

Anti-carbamylated protein antibodies are associated with long term disability and increased disease activity in patients with early inflammatory arthritis: results from the Norfolk Arthritis Register (NOAR)

Jennifer H Humphreys[1] , Marije K Verheul[2], Anne Barton[4,5], Alexander J MacGregor[3], Mark Lunt[1]
Rene E M Toes[2], Deborah PM Symmons[1,5], Leendert A Trouw[2], Suzanne MM Verstappen[1]

[1]Arthritis Research UK Centre for Epidemiology, University of Manchester, Manchester
[2]Leiden University Medical Center, Department of Rheumatology, Leiden, The Netherlands
[3]Arthritis Research UK Centre for Genetics and Genomics, University of Manchester, Manchester
[4]Norfolk and Norwich University Hospital, Norwich, United Kingdom
[5] NIHR Manchester Musculoskeletal Biomedical Research Unit, Manchester Academic Health Science Centre, Manchester, UK

Word count: 3045 (including abstract)

Key words: Early arthritis, rheumatoid arthritis, epidemiology, autoantibodies, disability, disease activity

Correspondence to: Dr Suzanne M.M. Verstappen, Arthritis Research UK Centre for Epidemiology, Manchester Academic Health Science Centre, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, UK

suzanne.verstappen@manchester.ac.uk

Tel: +44 (0)161 275 5663

Abstract

Objectives: Anti-carbamylated protein (anti-CarP) antibodies are a novel family of autoantibodies recently identified in patients with inflammatory arthritis. The aim of this study was to investigate their association with long term outcomes of disability and disease activity over 20 years follow up in a cohort of patients with inflammatory polyarthritis (IP).

Methods: NOAR recruited adults with recent onset swelling of ≥2 joints for ≥4 weeks from 1990 until 2009. At baseline HAQ and DAS28 scores were obtained and CRP, rheumatoid factor (RF), anticitrullinated protein antibodies (ACPA) and anti-CarP antibodies were measured. Further HAQ scores and DAS28 were obtained at regular intervals over 20 years. Generalized estimating equations were used to test the association between anti-CarP antibody status and longitudinal HAQ and DAS28 scores; then adjusting for age, gender, smoking status, year of inclusion and ACPA status. Analyses were repeated in subgroups stratified by ACPA status. The relative association of RF, ACPA and anti-CarP antibodies with HAQ and DAS28 scores was investigated using a random effects model.

Results: 1995 patients were included; 1310 (66%) were female. Anti-CarP antibodies were significantly associated with more disability and higher disease activity, HAQ multivariate β-coefficient (95% confidence interval) 0.12 (0.02-0.21), and these associations remained significant in the ACPA negative subgroups. The associations of RF, ACPA and anti-CarP antibodies were found to be additive in the random effects model.

Conclusion: Anti-CarP antibodies are associated with increased disability and higher disease activity in patients with IP. Our results suggest that measurement of anti-CarP antibodies may be useful in identifying ACPA negative patients with worse long term outcomes. Further, anti-CarP antibody status provided additional information to RF and ACPA.

Background:

Rheumatoid arthritis (RA) is a heterogenous inflammatory arthritis and individual patient outcomes can vary from mild to disabling and life limiting (1;2). The presence or absence of autoantibodies provides important prognostic information to clinicians and patients. Rheumatoid factor (RF) and, in particular, anti-citrullinated protein antibodies (ACPA) have been associated with more severe disease activity (3;4), greater levels of disability (5) and increased mortality (6). They also form part of the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria for RA (7). These criteria were developed with the aim of identifying patients with RA early in the natural history of the disease, using the initiating of disease modifying anti-rheumatic drugs (DMARDs) as their gold standard. Patients who lack ACPA and RF have been shown to be less likely to fulfil the 2010 RA criteria, although they may fulfil the older 1987 criteria (8;9). Nevertheless, in clinical practice there remains a subset of apparently seronegative patients who go on to experience high levels of disease activity and disability. If these patients could be distinguished from those patients with a milder disease course, they could benefit from early aggressive intervention.

Recently a new group of autoantibodies, anti-carbamylated protein (anti-CarP) antibodies, has been identified in the sera of patients with RA (10). These antibodies are directed against a post-translational modification of the amino acid lysine to homocitrulline in the presence of cyanate (11). They have been shown to pre-date the onset of symptoms (12-14) and may occur before or after the development of ACPA (12). Further, they have been shown to predict development of arthritis in patients with arthralgia (15). However, it is not yet known if they are associated with long term disability and disease activity. In addition, it would be clinically relevant to understand the influence of anti-CarP antibody status in patients with and without the other autoantibodies (RF and ACPA), as well as how much prognostic information is contributed by each antibody.

As patients with anti-CarP antibodies may lack RF or ACPA and therefore be less likely to fulfil RA criteria, it is important to study a broad group of patients presenting with inflammatory polyarthritis (IP) which would include a subgroup who meet RA criteria. The aims of this study were (i) to describe the prevalence and co-occurrence of RF, ACPA and anti-CarP antibodies in patients with IP, (ii) to investigate the relationship between anti-CarP antibody status and both disability and disease activity measured over time in patients presenting with IP, (iii) to investigate these relationships in ACPA positive and negative

subgroups, and (iv) to describe the additional predictive information provided by measuring these antibodies.

Methods:

Patients and follow up: Patients were included from the Norfolk Arthritis Register (NOAR). This cohort has been described previously (16). Briefly, adults >16 years old with at least 2 swollen joints for at least 4 weeks in the former Norfolk Health Authority area were recruited between 1990 and 2009. Patients recruited from 1995 to 1999 were excluded from this study as they were not followed beyond 2 years. Patients were also excluded if no serum sample obtained within the first year after recruitment was available. The selection of patients for the analysis is shown in full in the supplementary data file. At baseline in NOAR, patients are assessed by a nurse who obtains demographic details, medication details and smoking history, and performs a 51 tender and swollen joint count. The patients complete the British version of the Health Assessment Questionnaire (HAQ) (18). Patients in NOAR are followed up yearly for the first 3 years then at 5, 7, 10, 15 and 20 years from baseline. Patients repeat the HAQ and the nurse assessment at each follow up. Blood samples are taken at baseline and every 5 years thereafter, stored frozen and subsequently tested for RF (latex test, cut-off for a positive result >40 iu/l), ACPA (Axis-Shield Diastat Anti-CCP kit), and C-reactive protein (CRP) in Manchester, UK. The cut-offs for a positive test were set according to the manufacturer's guidelines at >40 iu/l for RF, >5 iu/l for ACPA, and >5 mg/l for CRP. The 3 item disease activity score (DAS28-CRP) (17) is calculated at baseline and every 5 years, and the 2010 ACR/EULAR criteria are applied retrospectively using baseline data. In 2013-2014, stored sera were sent to Leiden University Medical Center, The Netherlands, in a blinded fashion for measurement of anti-CarP antibodies using an in-house ELISA based on carbamylated fetal calf serum (FCS) as described before (10). Briefly, nonmodified FCS and modified-FCS were coated on Nunc Maxisorp plates (Thermo Scientific) overnight. After washing and blocking the wells were incubated with serum. Bound human IgG was detected using rabbit anti-human IgG antibodies (Dako), then HRP-labelled goat anti-rabbit IgG antibody (Dako). Following final washings, HRP enzyme activity was visualised using ABTS (10). NOAR is approved by the Norwich Local Research Ethics Committee and all patients gave written consent.

Statistical analysis: Differences in baseline disability (measured by the HAQ) and disease activity (measured by DAS28) between anti-CarP antibody positive and negative patients were evaluated using the Kruskal-Wallis test. Generalised estimating equations (GEE) were used to assess the association between anti-CarP antibody status and HAQ and DAS28 measured over time, allowing for the inclusion of patients with incomplete follow up data. A time interaction term was included to investigate any potential change in the relationship between baseline anti-CarP antibody status and HAQ or DAS28 scores. Univariate and subsequently multivariate models were constructed adjusting for age, gender, smoking status (stratified as current, previous or never smoked), polynomials of disease duration (to better fit the outcome measures), year of recruitment to NOAR and ACPA status. The analyses were repeated in the ACPA positive and negative subgroups and in patients who did and did not meet the 2010 RA classification criteria at baseline, omitting the ACPA confounder variable. In addition, as DAS28 was only available every 5 years, sensitivity analyses were performed using swollen joint count as an alternative measure of disease activity over time.

The individual effects of RF, ACPA and anti-CarP antibodies were then investigated. For each of the two outcomes of interest, a random effects model was used to test the association with each antibody. A three-way interaction term was included to investigate potential interactions between the antibodies; the resulting β-coefficient for each antibody estimated the added effect of that antibody. The final model was also adjusted for age, gender, smoking status, disease duration and year of recruitment.

A proportion of patients had anti-CarP antibodies tested but had missing data on some of the baseline covariates in the model (ACPA, CRP and smoking status, see table 1). To account for this, missing data were imputed using multiple imputation with chained equations and a sensitivity analysis was performed in the imputed dataset. In addition, as DAS28 was only available every 5 years, sensitivity analyses were All analyses were performed using STATA 12 software package (Stata, College Station, Tx, USA).

Results:

A total of 1995 patients with IP were included, 1310 (66%) were female and median age at onset (interquartile range (IQR)) was 55 (43-66) years. 460 (23%) patients were anti-CarP antibody positive and 1221 (61%) fulfilled the 2010 ACR/EULAR classification criteria for RA

at baseline. The median follow up time (IQR) was 8 (5-12) years. A summary of the baseline characteristics is shown in table 1. Baseline characteristics of patients who fulfilled the 2010 RA criteria are shown in supplementary table S1. A total of 1476 patients were tested for all three antibodies of whom 297 (20%) were anti-CarP antibody positive, with 74 (5%) testing positive for only anti-CarP antibodies. The distribution of all antibody statuses is shown in figure 1.

Levels of disability at baseline differed between anti-CarP positive and negative patients, respective median HAQ (IQR) 1.125 (0.5-1.75) and 0.875 (0.25-1.5), p<0.001. There were also differences in baseline DAS28 scores, respective median DAS28 (IQR) in the anti-CarP antibody positive and negative groups were 4.23 (3.19-5.31) and 3.73 (2.80-4.63), p<0.001.

Table 1. Baseline demographic and disease characteristics

	Total cohort n=1995	**Patients with all antibodies tested n= 1476**	**Missing *n (% total cohort)***
Female *n (%)*	1310 (66)	983 (67)	0
Age at symptom onset (years) *median (IQR)*	55 (43-66)	54 (42-65)	0
Smoking status *n (%)*			
Never	706/1982 (36)	535 (36)	13 (1%)
Previous	793/1982 (40)	585 (40)	
Current	483/1982 (24)	350 (24)	
Disease duration (weeks) *median (IQR)*	33 (17-69)	34(17-70)	0
HAQ *median (IQR)*	0.875 (0.375-1.5)	0.75 (0.25-1.5)	23 (1%)
DAS28 *median (IQR)*	3.81 (2.88-4.82)	3.79(2.85-4.78)	362 (18%)
RF positive *n (%)*	658/1895 (35)	463 (31)	100 (5%)
ACPA positive *n (%)*	389/1487 (26)	385 (26)	508 (25%)
Anti-CarP antibody positive *n (%)*	460 (23)	297 (20)	0
CRP, (mg/L) *median (IQR)*	8.7 (2-19)	8 (2-18)	362 (18%)
Satisfy 2010 RA classification criteria* *n(%)*	1221 (61)	893 (61)	0
On DMARDs at baseline assessment *n(%)*	722 (36)	501 (34)	0

IQR, inter-quartile range; HAQ, Health Assessment Questionnaire; DAS28, 28 joint disease activity score; RF, rheumatoid factor; ACPA, anti-citrullinated protein antibodies; Anti-CarP, anti-carbamylated protein antibodies; CRP, C-reactive protein; RA, rheumatoid arthritis; DMARDs, disease modifying anti-rheumatic drugs
* at baseline

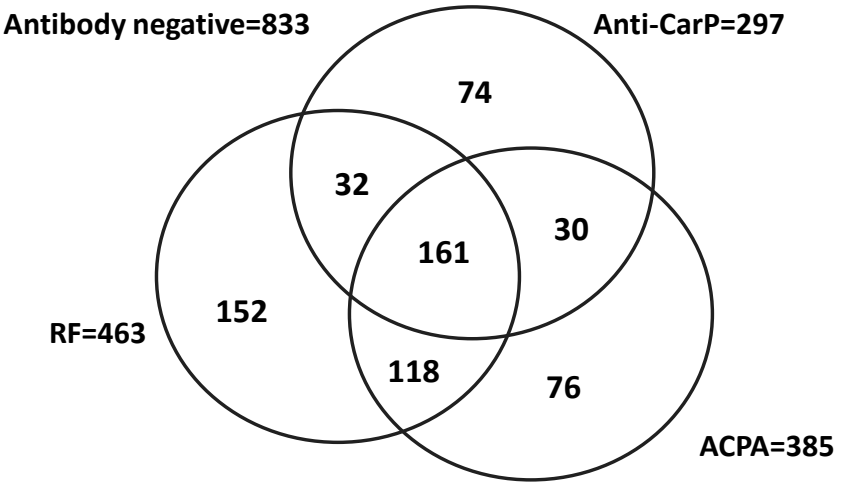


Figure 1. Distribution of antibodies in patients with IP who had all 3 antibodies tested

In the GEE model, patients who were anti-CarP antibody positive were, on average, more disabled at baseline and remained more disabled throughout follow up compared to those who were negative (figure 2), unadjusted GEE β-coefficient (95% CI) 0.21 (0.14,0.29), and this remained significant in the multivariate analysis including adjustment for ACPA status (table 2). Similarly, when DAS28 was the outcome of interest, anti-CarP antibody positive patients had, on average, higher levels of disease activity over time, unadjusted GEE β-coefficient 0.38 (0.25,0.50) (supplementary file figure S2). This association persisted in the multivariate model. In the ACPA negative subgroup there was also a significant association between anti-CarP antibody positivity and HAQ. It should be noted that for both HAQ and DAS28, the multivariate β-coefficient (95% CI) were very similar between the ACPA negative and ACPA positive groups and these estimates were not significantly different from each other. The interaction with time covariate was not statistically significant, meaning that the difference in HAQ scores between the average anti-CarP antibody positive patient and the average anti-CarP antibody negative patient did not increase or decrease over follow up; this is displayed in figure 2. A time interaction term was therefore not included in the final models.

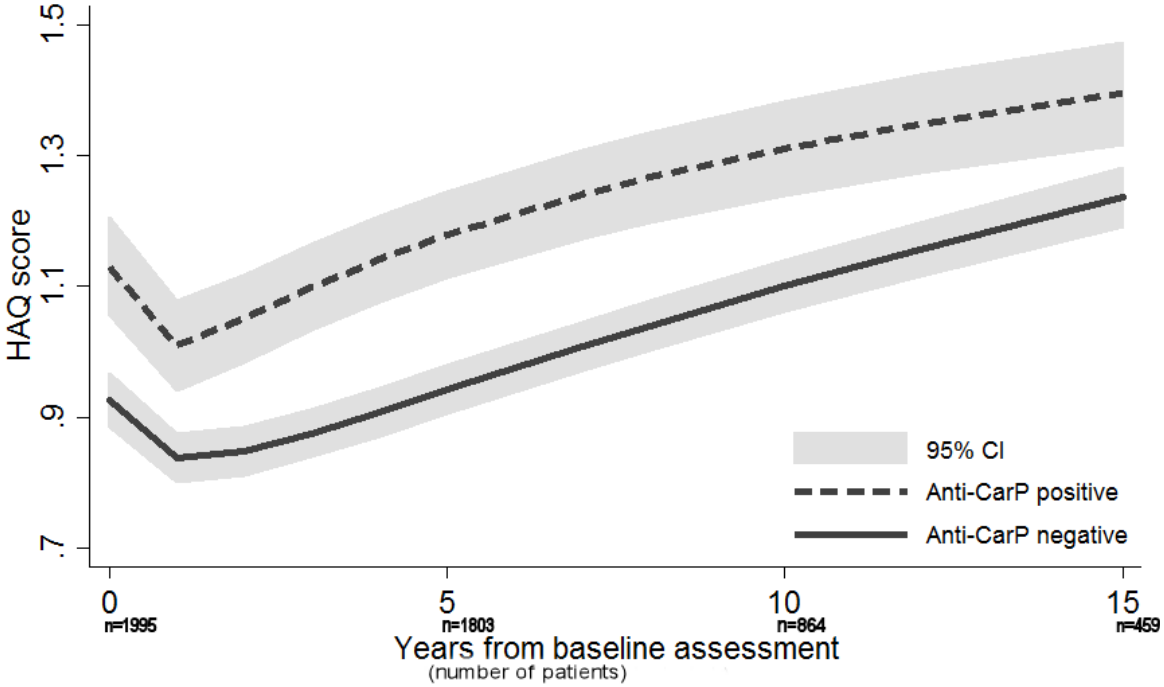


Figure 2 HAQ scores over time by anti-CarP antibody status (modelled by univariate GEE)

In patients who fulfilled the 2010 RA criteria at baseline, anti-CarP antibody status was associated with DAS28 and there was a trend to statistical significance with HAQ, respective multivariate β-coefficient (95% CI) 0.18 (0.04,0.32) and 0.07 (-0.01,0.16) (supplementary table S2). Interestingly there was a significant association with the HAQ amongst the group of patients who did not fulfil the 2010 RA criteria multivariate β-coefficient (95% CI) 0.19 (0.06,0.33). The sensitivity analysis with imputed missing covariates produced similar results (supplementary table S3), as did the sensitivity analysis with swollen joints as the outcome of interest (supplementary table S4).

Table 2. Association between anti-CarP antibody positivity and HAQ and DAS28

	Total cohort *β (95% CI)*	**ACPA +ve** *β (95% CI)*	**ACPA –ve** *β (95% CI)*
HAQ			
Univariate	0.21 (0.14,0.29) $	0.10 (-0.04,0.24)	0.18 (0.04,0.32) †
Multivariate*	0.12 (0.02,0.21) †	0.09 (-0.05,0.23)	0.14 (0.01,0.27) †
DAS28			
Univariate	0.38 (0.26,0.50) $	0.23 (0.01,0.46) †	0.11 (-0.11,0.34)
Multivariate*	0.23 (0.07,0.39) †	0.25 (0.03,0.48) †	0.18 (-0.03,0.40)

+ve:positive, -ve:negative

*adjusted for age, gender, smoking status, polynomials of disease duration, and year of recruitment

†p<0.05

$p<0.001

In the model which assessed the relative contributions of ACPA, RF and anti-CarP antibodies to long term disability, no interaction was found between the antibodies, and the R-squared value of the models were very similar with and without the interaction term. Therefore the effect of each antibody could be considered to be additive rather than multiplicative. Both ACPA and anti-CarP antibodies were significantly associated with long term disability, as measured by the HAQ, and had similar effect sizes, respective adjusted β-coefficient (95% CI) 0.12 (0.02-0.21) and 0.13 (0.03-0.21) (table 3). However, in the adjusted model RF was not. In terms of disease activity over time, again ACPA and anti-CarP antibodies were significantly associated with DAS28 score over time and RF was not (table 3).

Table 3. Association between autoantibodies and HAQ and DAS28

	Univariate *β (95% CI)*	**Multivariate*** *β (95% CI)*
HAQ		
ACPA	0.20 (0.12,0.28) $	0.12 (0.02,0.21) †
RF	0.12 (0.05,0.18) $	-0.03 (-0.12,0.05)
Anti-CarP antibodies	0.21 (0.14,0.29) $	0.13 (0.03,0.21) †
DAS28		
ACPA	0.36 (0.23,0.50) $	0.26 (0.09,0.43) $
RF	0.28 (0.17,0.39) $	-0.01 (-0.17,0.15)
Anti-CarP antibodies	0.38 (0.26,0.50) $	0.25 (0.09,0.42) $

*adjusted for age, gender, smoking status, polynomials of disease duration, and year of recruitment

†p<0.05
$p<0.01

Discussion:

This is the first study to investigate the associations between anti-CarP antibody status and long term disease activity and disability in patients with IP. We have shown that patients with anti-CarP antibodies are more disabled and have higher disease activity early in the disease and continue to have more functional disability and disease activity, compared to anti-CarP antibody negative patients. We have also shown that the influence of anti-CarP antibody positivity is similar to that of ACPA when considering these outcomes, and that measurement of the different autoantibodies provides additional information.

The majority of anti-CarP antibody positive patients in our study also demonstrated the presence of another antibody; however there was a subset of patients who were only positive for anti-CarP antibodies. Of particular interest are the associations with poor

outcomes in the ACPA negative subgroup, and the model adjusting for ACPA status. In general, ACPA negative patients are considered to have a good prognosis (3). However there is a small group who do poorly. For example, in studies of early arthritis cohorts, most patients who only fulfil the 1987 classification criteria for RA (characterised by the hallmarks of established RA such as radiological damage and nodulosis) and not the 2010 criteria (characterised by raised inflammatory markers and swollen/tender joint counts) are negative for RF and ACPA (8;9;19). Knowledge of anti-CarP antibody status in these patients therefore may be especially useful. In line with our results, other cohorts have demonstrated an association between anti-CarP antibody positivity and greater radiographic damage in patients with inflammatory arthritis and the subgroup of these who are ACPA negative (10;14). In our study, as well as stratifying patients with IP into ACPA positive and negative subgroups, we also adjusted for ACPA in the analyses of the whole cohort. This is because a number of studies have demonstrated that multiple autoantibodies can be accumulated in the preclinical phase of rheumatoid arthritis (20;21), possibly via the mechanism of epitope spreading, and ACPA usually appears before RF (22). It therefore seemed reasonable to consider baseline ACPA status a potential confounder.

In this study, we have addressed, for the first time, the 'added value' of testing for anti-CarP antibodies. Recent studies in the literature have investigated the influence of the number of autoantibodies on disease outcomes (6;23). Therefore, in addition to investigating the independent association of anti-CarP antibodies with disease outcomes, we wanted to address the additional information gained when RF and ACPA status is already known. It was interesting to note in this analysis that the coefficients for ACPA and anti-CarP antibodies were very similar. This suggests that, in terms of disability and disease activity over time, the impact of ACPA and anti-CarP antibodies are similar in patients with IP who test positive for these antibodies. Given our results demonstrated an additive effect of each antibody, it may therefore be useful to test more than one antibody in clinical practice when trying to assess current and future disability and disease activity.

There are some limitations in our study. There are currently no commercial assays available to test for anti-CarP antibodies, which could restrict the clinical impact of these results. However, the assay based on the methods described by Shi et al (10) has begun to be used more widely; to date it has been employed by two independent groups (24;25). In addition, a number of companies are developing routine assays to measure anti-CarP antibodies which should become available in the near future.

It is important to acknowledge that the effect sizes demonstrated in this study do not meet some previously published 'minimum clinically important difference' (MCID) for the HAQ (26). However these MCIDs were calculated and validated for use in clinical trials to test the effect of a specific treatment over a set period of time. Others have argued that the MCID estimates may be as low as 0.09 in observational studies (27). Our results certainly exceed this threshold. As mentioned above, the association of ACPA status with both HAQ and DAS28 demonstrated similar effect sizes to anti-CarP antibody status.

A larger proportion of patients recruited into NOAR were negative for all autoantibodies tested. This reflects the fact that patients with IP are a broad group which includes a subset of patients with RA, and that the majority of patients are presenting early in their disease course. We have previously shown in this cohort that 75-95% of patients recruited go on to satisfy the 1987 RA criteria (28). However it is also important to note that the 2010 RA criteria do not identify all patients with inflammatory arthritis who subsequently have poor outcomes, and this is particularly seen in seronegative patients (29). It was interesting therefore that anti-CarP antibody positivity was associated with significantly higher HAQ scores in the subgroup of patients who did not satisfy the 2010 criteria at baseline. In these patients, anti-CarP antibodies may be a marker of those who will go on to develop RA.

We have not taken into account treatment in our analysis. However we did include the year of registration in the multivariate models, which would account for changes in prescribing patterns since 1990. Importantly anti-CarP antibodies and ACPA were tested on stored sera; therefore the results were not known to the treating clinicians and could not have influenced treatment decisions. ACPA status may have been available through testing in routine clinical practice; however this would only apply to a small sample of NOAR patients seen by rheumatologists since 2009 when the test became widely available in Norfolk. In addition, as the anti-CarP antibody ELISA is a relatively new test it is not yet clear whether prolonged storage of sera before testing may influence the results; adjustment for year of registration to NOAR will have taken some of this effect into account. The anti-CarP antibody positive patients had more active disease and more disability at baseline and thus may have had more intensive therapy, potentially introducing channelling bias. However, by not including the impact of treatment we have biased our results towards the null hypothesis, and they are therefore likely to be an underestimate in terms of statistical significance. A further limitation is that we were not able to test the association between anti-CarP antibody status and radiological damage over time; this is due to the fact that not all patients in NOAR have radiographs and, in those that do, they

are taken at different follow ups depending on when the patient was recruited into the cohort. As a result we would not be able to accurately describe the relationship over time. Finally, there were some missing data on covariates in this study; multiple imputation with chained equations was used to allow inclusion of the whole sample in a sensitivity analysis which gave similar results to the main findings.

This analysis has shown that anti-CarP antibodies may be an important additional family of antibodies in predicting long term outcomes in patients with inflammatory polyarthritis, and may useful to test in addition to ACPA and RF.

Acknowledgments: This collaboration was supported by IMI JU funded project BeTheCure, contract no. 115142-2 and the NIHR Manchester Musculoskeletal Biomedical Research Unit. The authors gratefully acknowledge the support of clinical staff in Norfolk, the NOAR research nurses and the NOAR patients. NOAR is funded by Arthritis Research UK grant reference numbers 20380 and 20385. JH Humphreys is funded by Arthritis Research UK grant reference number 19743.

Reference List

(1) van der Helm-van Mil AH, Breedveld FC, Huizinga TW. Aspects of early arthritis. Definition of disease states in early arthritis: remission versus minimal disease activity. Arthritis Res Ther 2006;8(4):216.

(2) Humphreys JH, Warner A, Chipping J, Marshall T, Lunt M, Symmons DP, et al. Mortality trends in patients with early rheumatoid arthritis over 20 years: results from the Norfolk Arthritis Register. Arthritis Care Res (Hoboken) 2014 Sep;66(9):1296-301.

(3) Quinn MA, Gough AK, Green MJ, Devlin J, Hensor EM, Greenstein A, et al. Anti-CCP antibodies measured at disease onset help identify seronegative rheumatoid arthritis and predict radiological and functional outcome. Rheumatology (Oxford) 2006 Apr;45(4):478-80.

(4) De RL, Peene I, Hoffman IE, Kruithof E, Union A, Meheus L, et al. Rheumatoid factor and anticitrullinated protein antibodies in rheumatoid arthritis: diagnostic value, associations with radiological progression rate, and extra-articular manifestations. Ann Rheum Dis 2004 Dec;63(12):1587-93.

(5) Shidara K, Inoue E, Hoshi D, Sato E, Nakajima A, Momohara S, et al. Anti-cyclic citrullinated peptide antibody predicts functional disability in patients with rheumatoid arthritis in a large prospective observational cohort in Japan. Rheumatol Int 2012 Feb;32(2):361-6.

(6) Humphreys JH, van Nies J, Chipping J, Marshall T, van der Helm-van Mil A, Symmons DP, et al. Rheumatoid Factor and Anti-Citrullinated Protein Antibody positivity, but not level, are associated with increased mortality in patients with rheumatoid arthritis: results from two large independent cohorts. Arthritis Research and Therapy 2014;in press.

(7) Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010 Sep;69(9):1580-8.

(8) Humphreys JH, Symmons DP. Postpublication validation of the 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: where do we stand? Curr Opin Rheumatol 2013 Mar;25(2):157-63.

(9) Jung SJ, Lee SW, Ha YJ, Lee KH, Kang Y, Park MC, et al. Patients with early arthritis who fulfil the 1987 ACR classification criteria for rheumatoid arthritis but not the 2010 ACR/EULAR criteria. Ann Rheum Dis 2012 Jun;71(6):1097-8.

(10) Shi J, Knevel R, Suwannalai P, van der Linden MP, Janssen GM, van Veelen PA, et al. Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. Proc Natl Acad Sci U S A 2011 Oct 18;108(42):17372-7.

(11) Burska AN, Hunt L, Boissinot M, Strollo R, Ryan BJ, Vital E, et al. Autoantibodies to posttranslational modifications in rheumatoid arthritis. Mediators Inflamm 2014;2014:492873.

(12) Shi J, van de Stadt LA, Levarht EW, Huizinga TW, Hamann D, Van SD, et al. Anti-carbamylated protein (anti-CarP) antibodies precede the onset of rheumatoid arthritis. Ann Rheum Dis 2014 Apr;73(4):780-3.

(13) Gan RW, Trouw LA, Shi J, Toes RE, Huizinga TW, Demoruelle MK, et al. Anti-carbamylated Protein Antibodies Are Present Prior to Rheumatoid Arthritis and Are Associated with Its Future Diagnosis. J Rheumatol 2015 Jan 15.

(14) Brink M, Verheul MK, Ronnelid J, Berglin E, Holmdahl R, Toes RE, et al. Anti-carbamylated protein antibodies in the pre-symptomatic phase of rheumatoid arthritis, their relationship with multiple anti-citrulline peptide antibodies and association with radiological damage. Arthritis Res Ther 2015 Dec;17(1):536.

(15) Shi J, van de Stadt LA, Levarht EW, Huizinga TW, Toes RE, Trouw LA, et al. Anti-carbamylated protein antibodies are present in arthralgia patients and predict the development of rheumatoid arthritis. Arthritis Rheum 2013 Apr;65(4):911-5.

(16) Symmons DP, Silman AJ. Aspects of early arthritis. What determines the evolution of early undifferentiated arthritis and rheumatoid arthritis? An update from the Norfolk Arthritis Register. Arthritis Res Ther 2006;8(4):214.

(17) Wolfe F, Kleinheksel SM, Cathey MA, Hawley DJ, Spitz PW, Fries JF. The clinical value of the Stanford Health Assessment Questionnaire Functional Disability Index in patients with rheumatoid arthritis. J Rheumatol 1988 Oct;15(10):1480-8.

(18) Department of Rheumatology UMCNN. DAS28. http://www.das-score.nl/ 2004. Available from: URL: http://www.das-score.nl/

(19 van der Linden MP, Knevel R, Huizinga TW, van der Helm-van Mil AH. Classification of rheumatoid arthritis: comparison of the 1987 American College of Rheumatology criteria and the 2010 American College of Rheumatology/European League Against Rheumatism criteria. Arthritis Rheum 2011 Jan;63(1):37-42.

(20) Arend WP, Firestein GS. Pre-rheumatoid arthritis: predisposition and transition to clinical synovitis. Nat Rev Rheumatol 2012 Oct;8(10):573-86.

(21) Sokolove J, Bromberg R, Deane KD, Lahey LJ, Derber LA, Chandra PE, et al. Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis. PLoS ONE 2012;7(5):e35296.

(22) Nielen MM, Van SD, Reesink HW, Van De Stadt RJ, van der Horst-Bruinsma IE, de Koning MH, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004 Feb;50(2):380-6.

(23) Hecht C, Englbrecht M, Rech J, Schmidt S, Araujo E, Engelke K, et al. Additive effect of anti-citrullinated protein antibodies and rheumatoid factor on bone erosions in patients with RA. Ann Rheum Dis 2014 Aug 12.

(24) Yee A, Webb T, Seaman A, Infantino M, Meacci F, Manfredi M, et al. Anti-CarP antibodies as promising marker to measure joint damage and disease activity in patients with rheumatoid arthritis. Immunol Res 2015 Feb;61(1-2):24-30.

(25) Montes A, Perez-Pamoin E, Boveda MD, Gomez-Reino JJ, Gonzalez A. Antibodies against carbamylated proteins (anti-CarP), a new type of autoantibody, in spanish patients with rheumatoid arthritis. Annals of the Rheumatic Diseases 73[(Suppl2)], 129. 2014.
Ref Type: Abstract

(26) Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE, Jr. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis Rheum 2000 Jul;43(7):1478-87.

(27) Pope JE, Khanna D, Norrie D, Ouimet JM. The minimally important difference for the health assessment questionnaire in rheumatoid arthritis clinical practice is smaller than in randomized controlled trials. J Rheumatol 2009 Feb;36(2):254-9.

(28) Wiles N, Symmons DP, Harrison B, Barrett E, Barrett JH, Scott DG, et al. Estimating the incidence of rheumatoid arthritis: trying to hit a moving target? Arthritis Rheum 1999 Jul;42(7):1339-46.

(29) Burgers LE, van Nies JA, Ho LY, de Rooy DP, Huizinga TW, van der Helm-van Mil AH. Long-term outcome of rheumatoid arthritis defined according to the 2010-classification criteria. Ann Rheum Dis 2014 Feb;73(2):428-32.

6.2.1 Supplementary results

6.2.2 Differences between patients with missing and non-missing

As in section 6.1, the largest proportion of missing data was in ACPA status. There were no major differences between the patients with missing and non-missing baseline ACPA results (Table 6.2.1). Univariate analyses testing these potential differences are reported in the main manuscript in section 6.2.

Table 6.2.1 Baseline characteristics of patients with and without ACPA status

	Missing ACPA *n=508*	Non-missing *n=1487*
Female	322(63)	988(66)
Age onset	57(46-69)	54(42-65)
Symptom duration *(weeks)*	33(18-66)	33(17-69)
HAQ score	1(0.5-1.625)	0.75(0.25-1.5)
DAS28	3.97(3.06-4.94)	3.79(2.86-4.78)
CRP	12(3-28)	8(2-18)
Smoking status:		
current	127(25)	356(24)
previous	205(41)	588(40)
never	169(34)	537(36)
Satisfy 2010 RA criteria	321(63)	900(61)
Follow up time *(years)*	8.4 (5.0-10.3)	8.5(5.0-14.8)

Categorical variables are reported as n(%). Continuous variables are reported as median (interquartile range)

6.2.3 Online supplementary material for results paper 6.2

The following data in figures 6.2.1- 6.2.5 and tables 6.2.2 - 6.2.4 are referred to in the main manuscript in section 6.2 and formed the online supplementary material submitted to accompany this paper.

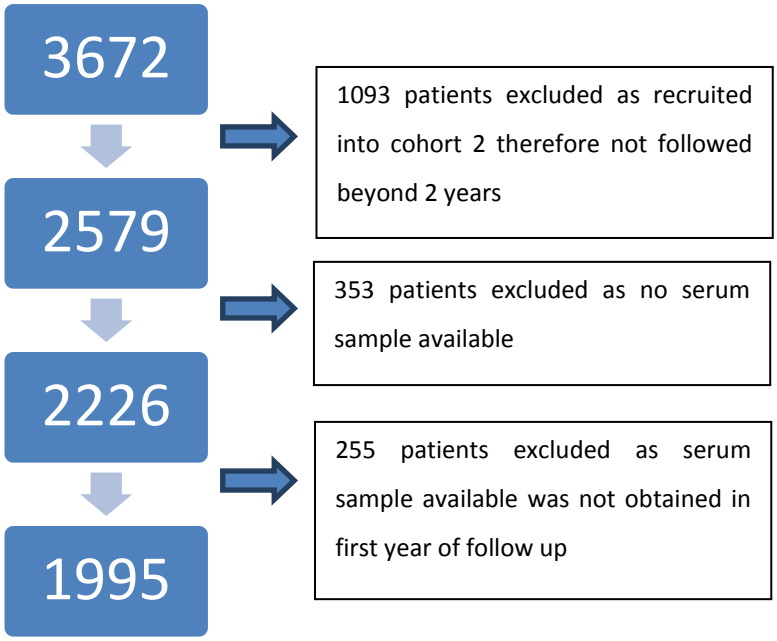


Figure 6.2.1 NOAR patient flow chart (supplementary figure S1)

Table 6.2.2 Baseline characteristics of IP patients who fulfilled 2010 RA criteria at baseline (supplementary table S1)

	RA* n=1221	Missing *n(% RA)*
Female *n (%)*	837(69)	0
Age at symptom onset (years) *median (IQR)*	55 (45-66)	0
Smoking status *n (%)*		7(1)
Never	408 (34)	
Previous	498 (41)	
Current	308 (25)	
Disease duration (weeks) *median (IQR)*	33 (18-67)	0
HAQ *median (IQR)*	1.125 (0.5-1.75)	14(1)
DAS28 *median (IQR)*	4.45 (3.69-5.37)	181(15)
RF positive *n (%)*	570 (48)	45(4)
ACPA positive *n (%)*	349 (39)	321(26)
Anti-CarP positive *n (%)*	372 (30)	0
CRP, (mg/L) *median (IQR)*	10.2 (3-23)	181(15)
On DMARDs at baseline assessment *n (%)*	502 (41)	0

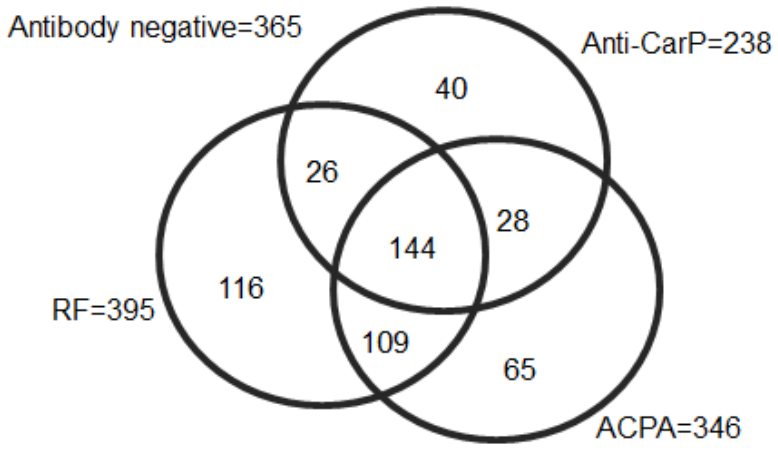


Figure 6.2.2 Distribution of antibodies in patients who fulfil 2010 RA criteria at baseline (supplementary figure S2)

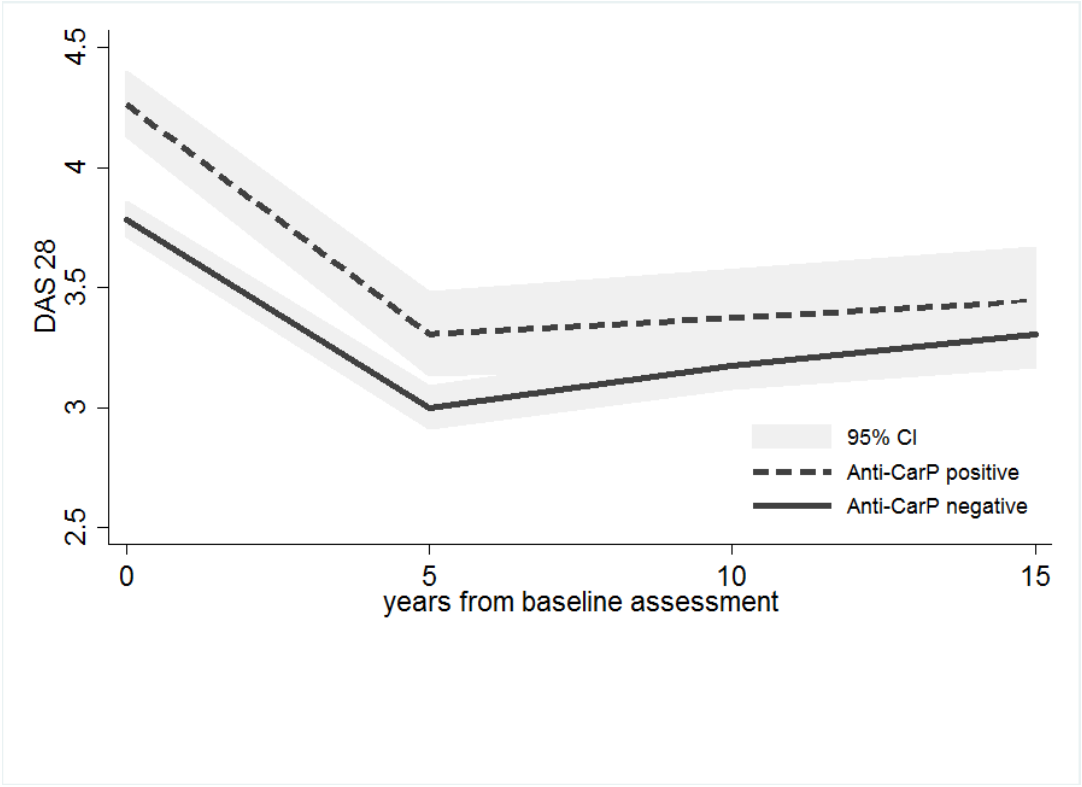


Figure 6.2.3 DAS28 scores over time by anti-CarP antibody status (modelled by univariate GEE) (supplementary figure S3)

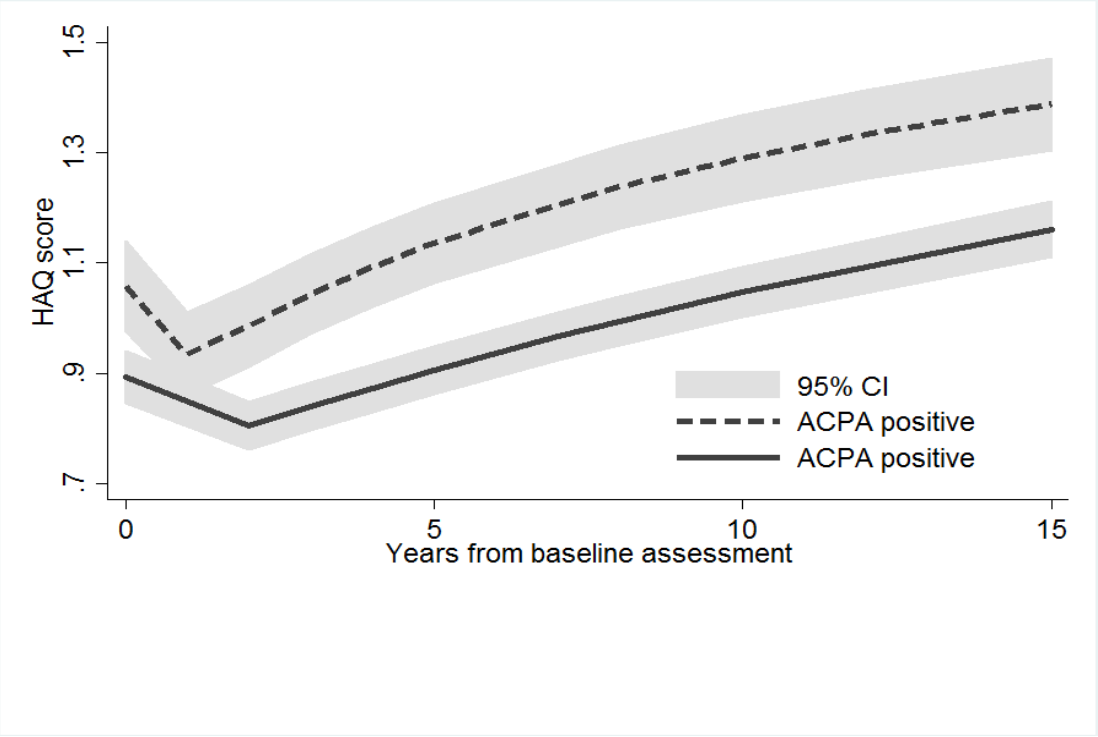


Figure 6.2.4 HAQ scores over time by ACPA status (modelled by univariate GEE) (supplementary figure S4)

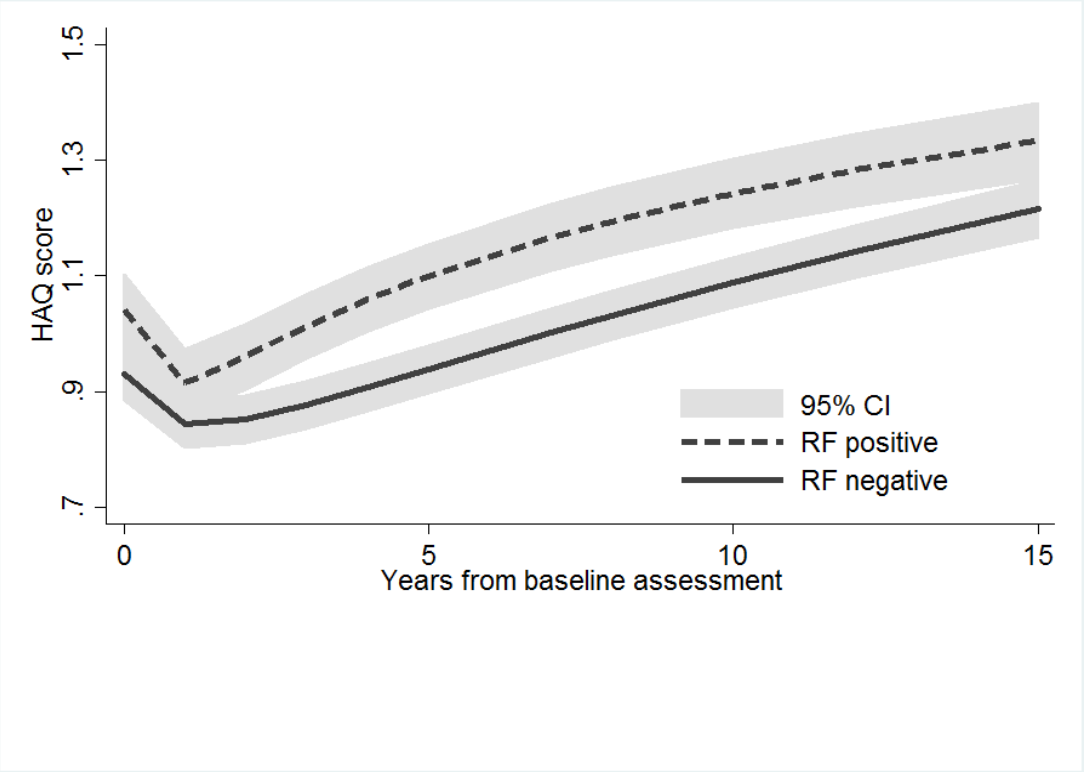


Figure 6.2.5 HAQ scores over time by RF status (modelled by univariate GEE) (supplementary figure S5)

Table 6.2.3 GEE models including all covariates, with subgroups of patients stratified by 2010 RA classification criteria and by ACPA status (supplementary table S2)

	Total cohort *β (95% CI)*	2010 RA +ve *β (95% CI)*	2010 RA -ve *β (95% CI)*	ACPA +ve *β (95% CI)*	ACPA –ve *β (95% CI)*
HAQ					
Anti-CarP	0.13 (0.03,0.23)	0.07 (-0.01,0.16)	0.19(0.06,0.33)	0.10 (-0.04,0.25)	0.15 (0.02-0.29)
Disease duration	0.03 (0.02,0.03)	0.03 (0.02,0.03)	0.02(0.02,0.02)	0.03 (0.03,0.04)	0.02 (0.02-0.03)
ACPA	0.10 (0.01,0.19)	-	-	-	-
Age	0.02 (0.01,0.02)	0.02 (0.01,0.02)	0.01(0.01,0.01)	0.02 (0.01,0.02)	0.01 (0.01-0.02)
Female gender	0.36 (0.29,0.44)	0.36 (0.27,0.45)	0.25(0.16,0.34)	0.26 (0.11,0.42)	0.40 (0.31-0.49)
Smoking *never*	ref	ref	ref	ref	ref
ex	0.04 (-0.04,0.13)	0.10 (0.00,0.19)	-0.04(-0.13,0.06)	-0.07 (-0.26,0.11)	0.08 (-0.01-0.17)
current	0.20 (0.10,0.30)	0.20 (0.09,0.31)	0.07(-0.04,0.19)	0.17 (-0.04,0.37)	0.20 (0.10-0.32)
Year of registration	0.01 (0.00,0.01)	0.00 (0.00,0.01)	0.01(0.00,0.02)	0.00 (-0.02,0.01)	0.01 (0.01-0.02)
DAS28					
Anti-CarP	0.23 (0.07,0.39)	0.18 (0.04,0.32)	0.01(-0.19,0.21)	0.25 (0.02,0.48)	0.22 (0.00-0.45)
Disease duration	-0.05 (-0.06,-0.04)	-0.09 (-0.10, -0.07)	0.01(0.00,0.02)	-0.06 (-0.08, -0.04)	-0.05 (-0.06- -0.04)
ACPA	0.27 (0.12,0.43)	-	-	-	-
Age	0.01 (0.00,0.01)	0.00 (0.00,0.01)	0.00(0.00,0.01)	0.00 (-0.01,0.01)	0.01 (0.00-0.01)
Female gender	0.49 (0.36,0.63)	0.26 (0.11,0.40)	0.32(0.17,0.46)	0.48 (0.23,0.72)	0.50 (0.35-0.66)
Smoking *never*	ref	ref	ref	ref	ref
ex	0.13 (-0.04,0.27)	0.06 (-0.10,0.21)	-0.02(-0.18,0.13)	0.03 (-0.26,0.32)	0.16 (-0.01-0.32)
current	0.16 (-0.01,0.32)	0.10 (-0.07,0.27)	0.02(-0.16,0.19)	0.12 (-0.2,0.44)	0.17 (-0.02-0.36)
Year of registration	0.00 (-0.01,0.01)	-0.01 (-0.02,0.00)	0.02(0.01,0.03)	-0.03 (-0.05, -0.01)	0.02 (0.00-0.03)

2010 RA +ve, satisfied the 2010 ACR/EULAR classification criteria for RA at baseline assessment; 2010 RA -ve, did not satisfy the 2010 ACR/EULAR classification criteria for RA at baseline assessment; HAQ health assessment questionnaire; Anti-CarP, anti-carbamylated protein antibodies; ACPA, anti-citrullinated protein antibodies; DAS28, disease activity score

Table 6.2.4 Sensitivity analysis performed with imputed dataset* (supplementary table S2)

	Total cohort *β (95% CI)*	2010 RA +ve *β (95% CI)*	2010 RA -ve *β (95% CI)*
HAQ			
Anti-CarP	0.16(0.09,0.24)	0.06(-0.03,0.14)	0.19(0.05,0.32)
Age	0.01(0.01,0.02)	0.01(0.01,0.02)	0.01(0.01,0.01)
Sex	0.35(0.29,0.41)	0.34(0.26,0.42)	0.26(0.18,0.35)
Disease duration	-0.07(-0.09,-0.05)	-0.10(0.13,-0.07)	-0.02(-0.05,0.01)
Disease duration^2	0.02(0.01,0.02)	0.02(0.02,0.03)	0.008(0.001,0.015)
Disease duration^3	-0.001(-0.001,-0.001)	-0.001(-0.002,-0.001)	0.000(-0.001,0.000)
Smoking *never*	Ref	Ref	ref
current	0.12(0.05,0.19)	0.14(0.04,0.23)	0.05(-0.04,0.14)
previous	0.03(-0.02,0.09)	0.05(-0.03,0.13)	-0.01(-0.09,0.07)
Year of registration	0.007(0.002,0.012)	0.005(-0.001,0.011)	0.01(0.004,0.018)
ACPA	0.03(-0.02,0.07)	0.006(-0.053,0.065)	0.05(-0.02,0.13)
DAS28			
Anti-CarP	0.29 (0.18,0.40)	0.14(.02,0.27)	0.32(0.13,0.52)
Age	0.003(0.000,0.007)	0.002(-0.002,0.006)	0.002(-0.002,0.006)
Sex	0.38(0.29,0.47)	0.34(0.22,0.46)	0.23(0.12,0.35)
Disease duration	-0.62(-0.71,-0.54)	-0.85(-0.96,-0.74)	-0.28(-0.39,-0.17)
Disease duration^2	0.11(0.09,0.14)	0.15(0.12,0.17)	0.06(0.04,0.09)
Disease duration^3	-0.007(-0.009,-0.005)	-0.009(-0.012,-0.007)	-0.004(-0.007,-0.002)
Smoking *never*	Ref	Ref	Ref
current	0.16(0.03,0.28)	0.15(-0.006,0.314)	0.06(-0.08,0.21)
previous	0.08(-0.03,0.19)	0.08(-0.05,0.21)	0.01(-0.13,0.15)
Year of registration	0.02(0.01,0.03)	0.02(0.01,0.03)	0.02(0.01,0.03)
ACPA	0.16(0.03,0.28)	-0.004(-0.147,0.140)	0.12(-0.09,0.32)

HAQ health assessment questionnaire; DAS28 disease activity score with 28 joints; Anti-CarP, anti-carbamylated protein antibodies; ACPA, anti-citrullinated protein antibodies

*The imputed dataset was not used to create models in the ACPA subgroups due to the proportion of missing ACPA

6.2.4 Sensitivity analyses using random effects model

Table 6.2.5 show the results of the random effects models constructed as a sensitivity analysis for the GEE models in paper 6.2, testing the association between anti-CarP antibody status and HAQ and DAS28 scores over time. These show very similar estimates to those seen in the GEE models, thus adding to the robustness of the results reported in paper 6.2.

Table 6.2.5 Random effects model of the association between anti-CarP antibody status and HAQ and DAS28

	Total cohort *β (95% CI)*	ACPA +ve *β (95% CI)*	ACPA –ve *β (95% CI)*
HAQ			
Univariate	0.21(0.13-0.29)	0.10(-0.04-0.25)	0.18(0.03-0.32)
Multivariate*	0.11(0.02-0.21)	0.09(-0.05-0.23)	0.13(0.01-0.27)
DAS28			
Univariate	0.38(0.26-0.50)	0.23(0.01-0.46)	0.11(-0.11-0.34)
Multivariate*	0.23(0.07-0.39)	0.26(0.04-0.49)	0.19(-0.03-0.41)

+ve:positive, -ve:negative, HAQ health assessment questionnaire, DAS28 disease activity score with 28 joints, ACPA, anti-citrullinated protein antibodies

*adjusted for age, gender, smoking status, polynomials of disease duration, and year of recruitment

Chapter 7: Results

Applying the 2010 criteria in patients with established inflammatory arthritis

This chapter comprises one manuscript which summarises and discusses the challenges of trying to apply the 2010 criteria in patients with longstanding established inflammatory arthritis, based on the views of experts in the field.

7 Applying the 2010 criteria in established inflammatory arthritis

To fully address objective 7 of this thesis, a validation study investigating the application of the 2010 criteria in patients with established disease is required. However, the NOAR cohort is primarily comprised of patients with incident EIA. To attempt this kind of validation study within NOAR, the 2010 criteria could be applied cross-sectionally after patients had had their symptoms for a set period of time, for example 5 or 10 years. However there are problems with this approach. Firstly the time point at which to apply the criteria would be completely arbitrary as there is no definition of established disease. More importantly, patients in NOAR are under assessment from the onset of their symptoms or very soon after and therefore will have already provided large amounts of data to the register, in addition to what might have routinely been collected for medical records. As a result, this would not be good representation of the circumstances of trying to apply the 2010 criteria *de novo* in patients who have had their disease for some time but are not already part of a research study. The generalisability of this study would be limited and such a study would not provide useful validation of the criteria in this setting. Therefore no data analyses in NOAR were undertaken.

As a data driven validation study within NOAR would not address the objective, a pilot analysis was done using data collected for a prevalence study in Norfolk in 2000 (1). The original study was designed to classify patients using the 1987 criteria and as a result there were significant missing data and it was not possible to estimate RA prevalence defined by the 2010 criteria. Nevertheless, amongst those patients who did have data on all variables within the 2010 criteria and even with the inclusion of the x-ray criteria, a proportion of patients could only be classified as RA using the1987 criteria. This therefore highlighted that the 2010 criteria in its current format may not be appropriate for the classification of patients with longstanding inflammatory arthritis. Following this, to explore the question further, expert opinion was sought on this topic. The custodians of a number of key long term observational cohorts of inflammatory arthritis across Europe were approached to participate by an email introducing the question and highlighting some of the areas of potential discourse; this can be found in appendix 4. The email also included a short questionnaire which asked for their views on how best to validate and apply the 2010 ACR/EULAR

classification criteria for RA in patients with longstanding inflammatory arthritis. The responses were then collated and the ideas developed into a viewpoint article, which was then reviewed and amended by all respondents to the questionnaire who are co-authors on the paper. The aim was to raise this challenging issue within the research community, to promote discussion and encourage formal attempts to address it. This chapter comprises the resulting manuscript.

7.1 Viewpoint: How do we classify rheumatoid arthritis in established disease – Can we apply the 2010 ACR/EULAR classification criteria?

Humphreys JH, Verstappen SM, Scire CA, Uhlig T, Fautrel B, Sokka T, Symmons DP. *Journal of Rheumatology* 2014 Dec;41(12):2347-5

The following statement is included at the request of the journal:

Viewpoint: How do we classify rheumatoid arthritis in established disease – Can we apply the 2010 ACR/EULAR classification criteria?

JH Humphreys[1], SMM Verstappen[1], CA Scire[2], T Uhlig[3], B Fautrel B[4], T Sokka [5], DPM Symmons[1,6]

1. Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Institute of Inflammation and Repair, The University of Manchester, Manchester Academic Health Science Centre, United Kingdom
2. Epidemiology Unit, Italian Society for Rheumatology, Milan, Italy
3. National Resource Center for Rehabilitation in Rheumatology, Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway
4. Université Pierre et Marie Curie, Faculté de Médecine, Department of Rheumatology, APHP, Hôpital Pitié Salpétrière, Paris, France
5. Department of Rheumatology, Jyväskylä Central Hospital, Jyväskylä, Finland
6. NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust and University of Manchester partnership

Word count: 1719

Correspondence to: Professor Deborah PM Symmons, Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Institute of Inflammation and Repair, The University of Manchester, Manchester Academic Health Science Centre, United Kingdom.

deborah.symmons@manchester.ac.uk

Tel: +44 (0)161 275 5044

Editorial

The creation of the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria for rheumatoid arthritis (RA) (figure 1)[1] sought to address some of the criticisms of the previous ACR criteria set published in 1987 (figure 2)[2]– namely, that they were insensitive, particularly early in the disease course, and that this led to the exclusion of many patients with early disease from clinical trials and research studies[3]. As a result, there was an absence of evidence regarding the efficacy of new treatments in this group of patients, who may have had the potential to benefit the most[4,5]. Thus one aim of the 2010 criteria was to identify those patients with early RA with the key purpose of rapid disease modifying anti-rheumatic drug (DMARD) initiation. However, the ability to classify patients as having RA is important at all phases of the disease, early and late. A case definition is required as an entry criterion not only to clinical trials, but also a consensus for inclusion in long-term observational studies and the whole spectrum of research in RA. Numerous studies have tested the validity of the new criteria since their publication against various standards[6], but they have generally been applied in patients with relatively short duration of symptoms, ranging from <3 months to <2 years. Thus the question remains, can we extend these classification criteria to patients with established disease?

Target population: Patients who (i) have at least one joint with clinical synovitis, and (ii) with the synovitis not better explained by another disease

	Score
A. Joint involvement (tender/swollen)	
1 large joint	0
2-10 large joints	1
1-3 small joints (+/- involvement of large joints)	2
4-10 small joints (+/- involvement of large joints)	3
>10 joints (at least 1 small joint)	5
B. Serology	
Negative RF&ACPA	0
Low-positive RF/low positive ACPA	2
High positive RF/high-positive ACPA	3
C. Acute-phase reactants	
Normal CRP&ESR	0
Abnormal CRP/ESR	1
D. Duration of symptoms	
<6 weeks	0
≥6 weeks	1

Add score of categories A-D:

≥6/10 = definite RA

Figure 1. 2010 ACR/EULAR Classification criteria for RA (1).
RF = rheumatoid factor; ACPA = anti-citrullinated peptide antibodies; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; ACR = American College of Rheumatology; EULAR = European League Against Rheumatism

Approach to date

The committee who developed the 2010 criteria emphasised they should encompass all patients that conform to the construct of 'rheumatoid arthritis'[1]. To this end, they recommended two approaches for patients with late disease. Firstly, they stated that the criteria could be applied retrospectively to those patients with the relevant available data[1]. There is inherent selection bias in this approach, which relies on adequate previous documentation unlikely to be available for all patients. In fact there may be less documentation in those whom the diagnosis was overt at initial presentation, than in patients who required investigation over time to make a diagnosis. Secondly, they recommended that patients with radiographic evidence of erosions typical of RA should be considered as having *prima facie* evidence of RA, and automatically be classified as such[1]. Van der Heijde et al established a definition of erosive disease as 'evidence of at least 3 eroded joints on X-rays of hands and feet which are available at the time of classification'[7,8]. However, these studies used baseline X-rays in two cohorts of patients with early inflammatory arthritis (symptom duration <2 years), the Dutch EAC and French ESPOIR cohort.

Criterion	Definition
Criteria 1-4 must have been present for ≥6 weeks	
1. Morning stiffness	Morning stiffness in and around the joints, lasting at least an hour before maximal improvement
2. Arthritis of ≥3 joints areas	≥3 joints areas simultaneously have had synovitis observed ay a physician
3. Arthritis of hand joints	At least 1 area swollen in a wrist, MCP or PIP joint
4. Symmetric arthritis	Simultaneous involvement of the same joint areas on both sides of the body
5. Rheumatoid nodules	Subcutaneous nodules, over bony prominences, extensor surfaces or juxta-articular regions
6. Serum rheumatoid factor (RF)	Positive RF
7. Radiographic changes	Radiographic changes typical of RA in posteroanterior hand and wrist radiographs
≥4/7 criteria satisfied = RA	

Figure 2. 1987 ACR Classification criteria for RA (2)

Challenges

The approaches so far give us no clues whether the 2010 criteria are useful and valid for patients with established disease. Indeed, there are considerable challenges to overcome if

we wish to validate the 2010 criteria in these patients. For example, application of the erosion criteria could potentially classify some patients with psoriatic arthritis, gout or erosive osteoarthritis as having RA, reducing the specificity of the criteria. These patients may fall under the exclusion criteria as having their synovitis better explained by another disease[1], however, since one purpose of classification criteria is to provide a homogenous group of patients for clinical trials, it has been suggested that any reduction in specificity seen with the 2010 criteria compared to the 1987 criteria is detrimental[8].

An additional challenge is raised by the fact that the 2010 criteria are based on active joint involvement and raised inflammatory markers (figure 1). There is thus potential for false negatives. Indeed, the 2010 criteria require *a priori* that a patient should have at least one joint with clinical synovitis at the time of applying the criteria, thus excluding any patient whose disease is in remission as a result of appropriate treatment. Given the current emphasis on aggressive management and aiming for remission, a proportion of patients with established disease may fail to satisfy the 2010 criteria if applied cross-sectionally, or if the relevant information (e.g. anti-cyclic citrullinated peptide (anti-CCP)) was not collected before they were started treatment.

It may be necessary to adapt or change the 2010 criteria for use in established disease. This leads to the question of how to define 'early' and 'established disease'; furthermore, at what point does the transition from one to the other occur? Is transition only based on symptom duration or also on clinical and radiographic features or even a change in patient reported outcomes?

Potential solutions

How could we address these challenges? Below are outlined some potential approaches:

- **Modification of the 2010 criteria**
- **Use of 1987 criteria in established RA**
- **Use of alternative criteria**

Modification of the 2010 criteria

Could we adapt the 2010 criteria for use in established disease? Adaptations of the 1987 ACR criteria have previously been evaluated as well, for example, to include X-rays of the

feet[9], or allowing the replacement of active joint inflammation with deformed joints if X-rays were not available[10].

Potential adaptations to the 2010 criteria might be the inclusion of some other typical RA features, for example rheumatoid nodules, radiographic erosions, symmetrical joint involvement, classic deformities or extra-articular manifestations. Intensive modern treatment strategies however, decrease the prevalence of these 'typical' RA features, which are frequently the consequence of long periods of uncontrolled disease activity that we aim to avoid in the modern era[5,11].

It could be possible to further clarify 'synovitis not better explained by another disease', by constructing a list of differential diagnoses which need to be ruled out before the classification criteria can be applied. This list may be easier to develop than in early arthritis, as disease specific features, such as the characteristic structural damage of psoriatic arthritis, have had time to develop.

Another solution might be to apply different weighting to the four parameters of the criteria in established disease. For example, re-weighting might increase the value given to anti-citrullinated protein antibodies (ACPA) positivity, as this is recognised to be more specific for RA than rheumatoid factor (RF)[12]. Alternatively, a different cut-off could be employed. As the 2010 criteria were developed to identify patients with early disease requiring DMARDs, a lower cut-off might be applied to patients already taking DMARDS, in whom disease activity will be lower than those with newly diagnosed RA.

Importantly, any approach to modify the criteria would necessitate a data-driven validation, raising the question of how to best define an 'RA' gold standard in such an analysis. Furthermore, any modification may likely change the sensitivity and specificity.

Continued use of the 1987 criteria

An alternative to modifying the 2010 criteria would be to continue to use the 1987 criteria in these patients. The 1987 criteria have been shown to have better sensitivity and specificity in patients with longstanding disease, compared to early disease [13] There are a number of issues with this approach; firstly it would be cumbersome to maintain two criteria sets, secondly it would introduce ambiguity to the research employing both, and

finally it would essentially create two parallel disease definitions that may not, in fact, be describing exactly the same disease. However, the NOAR study demonstrated there is reasonably good overlap between the two criteria sets, if the 2010 criteria are applied soon after disease onset and the 1987 criteria are applied cumulatively over 5 years follow up[14].

Alternative criteria - treatment

A further response could be to consider a rheumatologist diagnosis of RA alongside evidence of response to DMARDs as equivalent to RA classification in established disease. This is a pragmatic solution, but could lead to significant heterogeneity, as the majority of DMARDs are also used to treat other inflammatory arthritides, and diagnosis is dependent on the opinion of the individual clinician. When physician diagnosis was used as the gold standard in validation studies of the 2010 criteria in early arthritis, results varied extremely widely between cohorts, with sensitivities ranging from 58% -87%[15,16]. This viewpoint, however, is not meant to address diagnostic criteria, but to discuss the use of the 2010 classification criteria in established RA.

Alternative criteria - autoantibodies

Within the 2010 criteria, there is strong weighting given to the autoantibodies RF and ACPA. Synovitis of just one small joint for more than 6 weeks, with high levels of RF or ACPA, is sufficient to classify a patient as having RA. However, in studies that have made direct comparisons most patients who satisfy the 1987 criteria but not the 2010 criteria are seronegative[17-19]. As stated above, ACPA is more specific than RF, thus in the context of established disease it may be necessary to distinguish between RF positive and ACPA positive patients.

Could we therefore define all RA using auto-antibodies? It could be postulated that patients with so-called 'seronegative disease', eventually differentiate into non-RA phenotypes. Other autoantibodies, such as anti –mutated citrullinated vimentin (anti-MCV), which is an ACPA that is not detected by the most widely used test, anti-CCP2, and anti-carbamylated protein (anti-CarP) antibodies have been identified in anti-CCP negative patients and been associated with poor prognosis[20,21]. It is possible therefore, that inclusion of other antibodies in the criteria could account for all those patients with poor outcomes. However, the supportive evidence around these antibodies as prognostic markers, other than RF and ACPA, is currently insufficient. Further, it would be important

to confirm the stability over time of any such antibodies in order to use them to classify patients with established disease.

Conclusions

The 2010 ACR/EULAR classification criteria for RA represent a major step forward in rheumatology for patients in the early stages of RA and have already begun to inform the research agenda and our clinical practice[8,22]. Notably, it is different to the previous definition of the disease, characterised by damage. . By allowing us to identify early those with poor prognostic markers requiring rapid intervention, this new definition perhaps better captures the current construct of RA. However it is important to remember that patients with longstanding disease numerically outnumber those with early disease and we also require a robust method for their classification. We have summarised some of the many challenges in using the 2010 criteria in patients who are further into the natural history of the disease, and presented some of the potential options for dealing with them. Notably, all of the proposed solutions come with their own advantages and disadvantages. With this viewpoint article we wish to stimulate the discussion of how to classify patients with established RA.

Acknowledgements: The authors would like to acknowledge Dr Annette van der Helm-van Mil for her contribution.

References

1. Aletaha D, Neogi T, Silman AJ et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69:1580-1588.

2. Arnett FC, Edworthy SM, Bloch DA et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315-24.

3. Harrison BJ, Symmons DP, Barrett EM et al. The performance of the 1987 ARA classification criteria for rheumatoid arthritis in a population based cohort of patients with early inflammatory polyarthritis. American Rheumatism Association. *J Rheumatol* 1998;25:2324-30.

4. van Dongen H, van AJ, Lard LR et al. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. *Arthritis Rheum* 2007;56:1424-32.

5. Rantalaiho V, Korpela M, Laasonen L et al. Early combination disease-modifying antirheumatic drug therapy and tight disease control improve long-term radiologic outcome in patients with early rheumatoid arthritis: the 11-year results of the Finnish Rheumatoid Arthritis Combination Therapy trial. *Arthritis Res Ther* 2010;12:R122.

6. Humphreys JH, Symmons DP. Postpublication validation of the 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: where do we stand? *Curr Opin Rheumatol* 2013;25:157-63.

7. van der Heijde D, van der Helm-van Mil AH, Aletaha D et al. EULAR definition of erosive disease in light of the 2010 ACR/EULAR rheumatoid arthritis classification criteria. *Ann Rheum Dis* 2013;72:479-81.

8. Zeidler H. Systemic literature review of the performance of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: good news of debatable significance. *Ann Rheum Dis* 2013;72:e21.

9. Paimela L. The radiographic criterion in the 1987 revised criteria for rheumatoid arthritis. Reassessment in a prospective study of early disease. *Arthritis Rheum* 1992;35:255-58.

10. MacGregor AJ. Classification criteria for rheumatoid arthritis. *Baillieres Clin Rheumatol* 1995;9:287-304.

11. Lindqvist E, Jonsson K, Saxne T et al. Course of radiographic damage over 10 years in a cohort with early rheumatoid arthritis. *Ann Rheum Dis* 2003;62:611-16.

12. Avouac J, Gossec L, Dougados M. Diagnostic and predictive value of anti-cyclic citrullinated protein antibodies in rheumatoid arthritis: a systematic literature review. *Ann Rheum Dis* 2006;65:845-51.

13. Banal F, Dougados M, Combescure C et al. Sensitivity and specificity of the American College of Rheumatology 1987 criteria for the diagnosis of rheumatoid arthritis according to disease duration: a systematic literature review and meta-analysis. *Ann Rheum Dis* 2009;68:1184-91.

14. Humphreys JH, Verstappen SM, Hyrich KL 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;72:1315-20.

15. Cornec D, Varache S, Morvan J et al. Comparison of ACR 1987 and ACR/EULAR 2010 criteria for predicting a 10-year diagnosis of rheumatoid arthritis. *Joint Bone Spine* 2012;79:581-85.

16. Reneses S, Pestana L, Garcia A. Comparison of the 1987 ACR criteria and the 2010 ACR/EULAR criteria in an inception cohort of patients with recent-onset inflammatory polyarthritis. *Clin Exp Rheumatol* 2012;30:417-20.

17. Cader MZ, Filer A, Hazlehurst J et al. Performance of the 2010 ACR/EULAR criteria for rheumatoid arthritis: comparison with 1987 ACR criteria in a very early synovitis cohort. *Ann Rheum Dis* 2011;70:949-55.

18. Jung SJ, Lee SW, Ha YJ et al. Patients with early arthritis who fulfil the 1987 ACR classification criteria for rheumatoid arthritis but not the 2010 ACR/EULAR criteria. *Ann Rheum Dis* 2012;71:1097-98.

19. Burgers LE, van Nies JA, Ho LY et al. Long-term outcome of Rheumatoid Arthritis defined according to the 2010-classification criteria. *Ann Rheum Dis* 2013.

20. Shi J, Knevel R, Suwannalai P et al. Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. *Proc Natl Acad Sci U S A* 2011;108:17372-77.

21. Luime JJ, Colin EM, Hazes JM et al. Does anti-mutated citrullinated vimentin have additional value as a serological marker in the diagnostic and prognostic investigation of patients with rheumatoid arthritis? A systematic review. *Ann Rheum Dis* 2010;69:337-44.

22. Bykerk VP, Massarotti EM. The new ACR/EULAR classification criteria for RA: how are the new criteria performing in the clinic? *Rheumatology (Oxford)* 2012; 51 Suppl 6:vi10-vi15.

Chapter 8: Discussion

In this chapter the previous results chapters are summarised and reviewed in the context of current medical literature. The strengths and weaknesses of the work are assessed and the implications for research and clinical practice are discussed.

8 Discussion

The overarching aim of this thesis was to validate the 2010 criteria ACR/EULAR classification criteria for RA beyond standard measurement of sensitivity and specificity. The analyses in this thesis have demonstrated that patients with IP who satisfy the 2010 criteria have increased mortality (chapter 5), disability, disease activity and radiographic change (chapter 6) compared to those who do not; thus demonstrating good construct validity. Further, the criteria were employed for the first time to measure the incidence of RA (chapter 4) and to describe mortality trends over time (chapter 5). Finally the predictive properties of a novel family of autoantibodies (anti-CarP antibodies) were tested. In this chapter the results will be summarised, along with their contribution to the current literature. For the published manuscripts, this will focus on new literature since the time of their publication. In addition, the strengths and weakness of the thesis will be discussed. The clinical implications of the results and future directions for research in this area will be proposed.

8.1 Summary of the main findings

8.1.1 Incidence of RA

The publication of a new set of classification criteria provided a new definition of RA. As a result, estimates of disease frequency based on the previous criteria set were no longer valid. In chapter 5 of this thesis, the incidence of RA was estimated for the first time using the 2010 criteria. The overall IR for the year 1990 was 40 per 100 000 population; for women it was 55 per 100 000 and for men 25 per 100 000. All of these estimates were higher than those produced when applying the 1987 criteria to the same population. The analysis then examined age and sex specific IRs; the peak age of incidence for women was found to be in the 55-64 age group, which was younger than in men where incidence peaked amongst patients over 65 years old. The final part of the analysis examined the cumulative age and sex-specific IRs of the 2 criteria sets. The IRs of patients who satisfied the 2010 criteria at baseline were very similar to the 5 year cumulative IRs using the 1987 criteria. This indicates that patients can be classified as having RA earlier in their disease course by the 2010 criteria. However there was a small subgroup of patients who met the 1987 criteria and were never classified as RA by the 2010

criteria; these patients were almost all seronegative, suggesting seronegative patients with poor outcomes may be missed by the 2010 criteria.

To date this remains the only published study estimating the incidence of RA using the 2010 criteria. Since its publication, studies from Sweden, Italy and Taiwan have also published incidence estimates of RA in their respective countries (161-163). However these studies all used either national registers or administrative databases in which there are insufficient clinical data to apply classification criteria for case identification, and therefore other definitions were used. In Sweden, data linkage between three national registers was used to define their cases of incident RA between 2006 and 2009 (161). Interestingly, their estimates were remarkably similar to ours with overall incidence of 41 per 100 000 population (56 and 25 per 100 000 for women and men respectively). Although not based on classification criteria, the authors describe an unpublished validation study of their case definition which demonstrated it was 90% concordant with either 1987 or 2010 criteria (161). The study from Italy estimated the annual incidence of RA in 2011, and used an administrative database to identify their cases. In Italy, if patients have received a confirmed diagnosis of RA from a specialist, they qualify for reimbursement of medical costs; this was combined with a period of DMARD prescription to define RA in this study. Although a less robust definition than the Swedish study, nevertheless they also demonstrated a similar overall incidence of 35 per 100 000 population. In contrast, Kuo et al (163) found much lower incidence of RA in Taiwan at 16 per 100 000 population, also using an administrative claims database. This may be due genetic differences, as there is evidence of a lower frequency of RA in Chinese populations (164-166), but may also reflect different environmental and socioeconomic factors.

It should be noted that the analysis in this thesis was based on retrospective application of the 2010 criteria to patients with symptom onset in the year 1990. The similarities of our estimates to those described above in Sweden and Italy, which were based on patients presenting 15 to 20 years later, suggest that the incidence of RA is no longer declining as reported in the latter part of the 20th century (167) and may now have plateaued, at least in Western populations. This is in keeping with reports from the Rochester Epidemiology Project in the USA (168), which has been reporting on the incidence for RA in Rochester County, Minnesota for over 50 years. Further, the Global Burden of Disease 2010 study, which gathered data on prevalence and incidence of RA around the world, also found that the burden and prevalence of RA remained stable between 1990 and

2010 (169). However, both the Rochester project and the Global Burden of Disease Study used the 1987 criteria as case definition because the 2010 criteria had not yet been published at the time of data collection and the variables requires to apply the 2010 criteria retrospectively were not available. Therefore, a key investigation absent from the current literature is a study of the prevalence of RA defined by the 2010 criteria. This would require application of the 2010 criteria cross-sectionally in patients with established RA, and there are a number of inherent challenges in such a study. These have been discussed in chapter 7 and are further discussed below in section 8.1.4. However, this would establish a complete picture of disease occurrence as described by the 2010 criteria.

8.1.2 Mortality in RA

Mortality and the 2010 criteria

In results chapter 5 aspects of the increased mortality in patients with RA were explored. The first analysis in section 5.1 demonstrated that patients with EIA who fulfilled the 2010 criteria had an increased risk of death compared to those who did not (age and sex adjusted HR 1.35, 95% CI 1.13-1.64). This demonstrates that the 2010 criteria identify patients with EIA at risk of poor outcomes, strengthening the construct validity of the criteria.

Number of autoantibodies and mortality

From the literature review and the analysis in chapter 4, it was evident that positive autoantibody status plays an important part in satisfying the 2010 criteria. Therefore the next analysis, reported in section 5.2, investigated the association between early death and RF and ACPA, both in term of seropositive or seronegative status and the level of the antibodies as defined by the 2010 criteria. This work was done in collaboration with investigators from the Leiden EAC, and demonstrated no consistent association between the level of individual autoantibodies and survival. However in both cohorts, the presence of both RF and ACPA, compared to seronegative status, increased the risk of early death (NOAR multivariate HR 1.43, 95% CI 1.13-1.82), whereas single antibody positivity did not. This was an interesting finding, as it suggests that the two antibodies have at least an additive, if not a multiplicative effect on mortality, although we did not explore which in this particular analysis. The results suggest that in patients presenting with EIA, the number of positive antibodies may be of greater importance than the levels of those antibodies. In addition, the association between ACPA positivity and mortality that had previously been demonstrated in

NOAR (53) was replicated in the Leiden EAC. This was an important result to report as this association has not been widely investigated in the literature.

There has been increasing interest in whether the presence of multiple autoantibodies is important in the prognosis of RA. Hecht et al recently investigated this in a study of patients with RA from Germany using quantitative computer tomography imaging to identify and characterise erosions (170). They found that the presence of ACPA and RF together was associated with increased prevalence of erosions, whereas single antibody positivity was not. In addition, Sokolove et al, in a cohort of veterans in the USA, identified that double positive autoantibody status was associated with higher levels of disease activity and swollen joint counts compared to single autoantibody positive or seronegative status (171). Van der Linden et al investigated whether testing for more than one of anti-CCP, RF, and anti-modified citrullinated vimenten (another ACPA not detected by anti-CCP) antibodies is useful in predicting outcomes in patients presenting with undifferentiated arthritis (UA) and RA (55). In contrast to the studies by Hecht and Sokolove, they found that testing two antibodies compared to one antibody did not improve the PPV or NPV for RA development from UA, and did not increase the predictive accuracy to detect radiographic progression or DMARD free remission in patients with RA. However, in a separate analysis they found that the more positive antibodies that were present, the greater the joint destruction and lower the chances of DMARD free remission.

Mortality over time

The final paper in this chapter examined trends in mortality rates over time in patients with EIA and RA classified by the 2010 criteria. Overall age and sex-specific mortality rates were increased in patients with EIA and RA compared to the background Norfolk population with respective SMRs (95% CI) of 1.16 (1.04-1.29) and 1.22 (1.07-1.40). Although the crude mortality rates decreased over time, the MRRs showed no change in the mortality rates once changes in the background population had been accounted for; compared to mortality rates in cohort 1 (1990-94), the MRRs (95% CI) of cohorts 2 (1995-99) and 3 (2000-04) respectively were 1.13 (0.77-1.67) and 1.08 (0.7-1.68). Deaths were included if they occurred with the first 7 years of follow up; this was done so that patients recruited into the earliest cohort did not have apparently higher mortality rates due to a longer follow up period in which they could die. It is therefore possible that there are differences in the mortality rates over time, but that they occur after the first 7 years and therefore would not have been identified in this analysis. There is some support for this in the literature, for example in an RA inception cohort in the Netherlands

Radovits et al demonstrated excess mortality was only evident after 10 years of disease (172).

A recent study of patients with RA in Canada showed very similar results to ours; mortality in RA decreased over time but remained elevated compared to the general population (173). A meta-analysis by Dadoun et al also found no significant change in mortality rates after adjustment for changes in background population mortality in studies dating from 1955-1995 (174). However, the most recent analysis from the Rochester Epidemiology project which concluded follow up in 2007 suggested that, compared to the general population, cardiovascular mortality in RA may have increased (175). In contrast, data from the Early Rheumatoid Arthritis Study (ERAS) and the Early Rheumatoid Arthritis Network (ERAN) cohorts in the UK has shown that life expectancy has increased in patients with RA, and these increases are over and above those seen in the UK population (176). There are significant expectations that RA specific mortality rates should improve, in response not only to improved treatment strategies in RA, but also increased awareness of the excess burden of cardiovascular disease (CVD) in RA (177) leading to opportunities for primary prevention of CVD. However, the current data are not yet sufficiently robust or consistent to support this claim. It should be noted that the ERAS/ERAN study included patients recruited between 2002 and 2012, and therefore is the most modern cohort of patients in any of these mortality studies. Therefore it is possible that the impact of modern treatment strategies is a latent effect which is only now becoming evident. An alternate explanation is that the mortality risk in patients with RA differs between Europe and the USA and this may be an area that warrants further investigation.

8.1.3 Longitudinal outcomes and novel antibodies in RA

Longitudinal outcomes

In chapter 6, longitudinal outcomes of patients with IP other than mortality were investigated. The first analysis in section 6.1 addressed the question 'do patients with IP who satisfy the 2010 criteria at first presentation have worse outcomes in terms of disability, disease activity and radiographic damage over follow up?' and whether that association was different in patients recruited before and after the year 2000. The analysis demonstrated that patients who met the criteria at baseline assessment were more disabled, had higher levels of disease activity and more radiographic damage throughout follow up compared to those patients who did not meet the criteria. Surprisingly, there was little difference in the strength of

these associations in disease activity and disability before and after the year 2000, for example the multivariate GEE β (95%CI) for the HAQ were 0.38 (0.32-0.44) and 0.42 (0.33-0.50) in the pre-and post-2000 subgroups respectively. It might have been expected to have diminished in recent years, as biologic drugs and aggressive treatment strategies should theoretically reduce the impact of disease on these long term outcomes. However we did not find evidence of this in this analysis. This may be because an observational cohort such as NOAR reflects 'real life' treatment practices. For a wide variety of reasons, such as the presence of comorbidity and patient and clinician preference, 'real life' may not always follow the aggressive combination and biologic therapy treatment strategies that were so effective in clinical trials (67;73;75). There were differences in the radiological outcomes, with damage being more severe in the group recruited pre-2000. However over a third of the x-rays in the pre-2000 subgroup were taken at the 5th anniversary, whereas for the post-2000 subgroup, almost all were taken within the first year of observation; as a result the early subgroup will inevitably show more evidence of radiological damage. Overall these findings provided further evidence of good construct validity for the 2010 criteria.

Nevertheless, along with the data from chapter 4 on the incidence of RA, it is clear that prognosis is not universally good for patients who do not satisfy the 2010 criteria. In this analysis, after excluding patients who were on DMARDs at baseline, at each follow up there were a proportion of patients who were subsequently treated with DMARDs despite not satisfying the 2010 criteria. This would suggest that their arthritis did not resolve and there was sufficient concern of poor outcome by the treating rheumatologist to warrant intervention. For example at the 5 year follow up, of the 768 patients who had not satisfied the criteria at baseline, 187 (24%) had been treated with DMARDs. Even if the 2010 criteria were applied cumulatively to that point, 19% of patients who did not satisfy the criteria still received DMARDs. Similarly, at baseline and throughout follow up there was a subgroup of patients who did not satisfy the 2010 criteria but nevertheless experienced moderate to severe disability with HAQ scores greater than 1; at the 5 year follow up assessment this was 25% of the total 2010 criteria negative group. This again highlights the need for other biomarkers to identify these patients and differentiate them from those with a milder disease course.

Studies have previously examined changes in disease severity over time. In the Wichita Arthritis Cohort in the USA, Finckh et al examined radiographic outcomes and disability over 10 years of follow up in RA patients recruited between 1973 and

1999 (178). Unlike in the analysis in this thesis, they found there was evidence of diminishing disease severity over the decades, and in a detailed and methodologically thorough analysis they were able to attribute that to improved disease management. However, it may be that there was a starker contrast between treatment in 1999 compared to 1973 (when few DMARDs were available and those that were available were used much later in the course of the disease), than would be seen between the eras compared in this analysis (1990-2000 compared to 2000-2009). In a study from an early arthritis inception cohort in Nijmegen, Netherlands, Welsing et al also found that DAS28 scores over the first 5 years of follow up had diminished from 1985-2000 (179). However, in addition they noted decreases in baseline disease activity over time and increased baseline disability, a trend which has also been identified in NOAR (160). This may be due in part to more rapid initiation of DMARDs prior to assessment in recent years, although the trend was still seen in a sensitivity analysis of DMARD naïve patients, as well as earlier referral of milder disease due to increased awareness of the 'window of opportunity'.

Anti-carbamylated protein antibodies

The second analysis in section 6.2 investigated whether a novel autoantibody, anti-CarP antibody, was associated with long term disability and disease activity. The analysis showed that patients with IP who were anti-CarP antibody positive had higher levels of disability and disease activity at their baseline assessment and throughout up to 20 years of follow up compared to anti-CarP antibody negative patients, multivariate GEE β (95%CI) for HAQ and DAS28 0.13 (0.03-0.23) and 0.22 (0.06-0.37) respectively. In addition, the value of testing for each of the three antibodies (RF, ACPA and anti-CarP) in combination was examined; this analysis found that ACPA and anti-CarP antibody status provided additional information to each other and RF status, but RF status had no significant association with the either HAQ or DAS28 when ACPA and anti-CarP antibody statuses were taken into account. For practising clinicians it might therefore be most useful to test for ACPA and anti-CarP antibodies in combination to obtain the greatest information about a patient's long term outcomes.

There is a growing body of literature around the role of anti-CarP antibodies in inflammatory arthritis. As well as being identified in a group of patients with established RA and found to be associated with radiographic damage in those patients (61), two independent studies have now demonstrated the presence of anti-CarP antibodies prior to the development of joint symptoms. The first was in

healthy blood donors who went on to develop RA (59), and the second in members of the US armed forces who donate blood samples throughout their years in services and later developed RA (180). Anti-CarP antibodies have also been shown to predict the future development of RA in patients presenting with arthralgia (58). The data from this thesis therefore add to this relatively limited literature base by providing information on the impact of anti-CarP antibody status on subsequent disability and disease activity once arthritis has developed.

As discussed above in section 8.1.3, there remains concern about patients with IP who are 'missed' by the 2010 criteria but still have poor prognosis. This may in part be due to their lack of RF or ACPA, as highlighted by the results in section 4.1 of this thesis, where a proportion of patients, who were almost all seronegative, did not satisfy the 2010 criteria but did satisfy the 1987 criteria. The results from the analysis in section 6.2 of this thesis suggest anti-CarP antibodies may help identify some of these patients. Further, even in those patients who do demonstrate ACPA and/or RF, anti-CarP antibodies have an independent association with poor outcomes. If this is considered alongside the results from paper 5.2 and the literature discussed in section 8.1.2, there are appears to be a useful role in testing for more than one antibody in clinical practice, and that should potentially include anti-CarP antibodies. This would require the replication of these results in other cohorts and evidence that anti-CarP antibodies are associated with further important RA outcomes such as mortality.

It is worth considering whether patients with IP who lack ACPA and RF have a different disease from RA. Prior to the recognition of ACPA, patients were often categorised as seropositive or seronegative based on the presence or absence of RF (181;182) and in clinical practice may have been treated differently as a result (182). The advent of ACPA and the considerable overlap of ACPA and RF positivity, as well as growing evidence that outcomes may differ dependent on the presence of one, two or more antibodies as discussed in section 8.1.2 and paper 5.2, make the distinction between seropositive and seronegative disease less clear. Nevertheless researchers have also debated whether ACPA positive and negative RA are the same disease (183). Certainly the underlying genetic predisposition appears to differ considerably between these two groups (184). It could be postulated that anti-CarP antibody status might be useful in classifying an altogether separate form of inflammatory arthritis from RA. However, the presence of all three antibodies in some patients with IP would make differentiating these diseases more difficult.

8.1.4 RA classification criteria in established disease

Expert opinion was sought on how to approach the application and validation of the 2010 criteria in patients with longstanding inflammatory arthritis. The manuscript in chapter 7 summarises these opinions, and highlights three main challenges in doing so. First, since the radiographic criterion bypasses the exclusion criteria, it may allow misclassification of patients as having RA who in fact have other diseases such as PsA, gout or erosive OA. Secondly patients on treatment may fail to satisfy the criteria if applied cross-sectionally (as the treatment could reduce the number of swollen or tender joints and normalise inflammatory markers), or if there are insufficient data within the medical records to allow retrospective classification. Thirdly any attempt to adapt the 2010 criteria for use in established disease requires definition of when 'early' disease becomes 'established' disease, as this varies widely in the literature (96;145). A number of solutions were offered including continued use of the 1987 criteria in parallel to the 2010, an approach that has been taken in some studies to date (185), re-weighting of the four parameters within the criteria, and adopting the prescription of DMARDs as a proxy for classification. It was concluded that all the proposed solutions have potential advantages and disadvantages. To date there remain no published data on the performance of the 2010 criteria in patients with longstanding disease.

8.2 Strengths and weaknesses of the analyses

8.2.1 NOAR design and setting

The key strengths of these analyses lie in the design of NOAR. The recruitment of patients with IP, a broad category of inflammatory arthritis which would include subsets satisfying both the 1987 and 2010 criteria, was essential to achieve the objectives of this thesis. Along with this, there was a major effort undertaken to ensure that all patients presenting to healthcare for the first time with symptoms suggestive of inflammatory arthritis were registered to NOAR, at least at the outset of the register. This gave sufficient confidence of near complete capture of cases to undertake an IR estimation.

Further, the background population setting of Norfolk is ideal for longitudinal studies. Norfolk is renowned as a setting for epidemiological studies due to its

stable population; in addition NOAR recruited from the former Norwich Health Authority area for which there was single point of referral to secondary care. Although its boundaries do not identically match the county boundaries, they are a very close approximation, and this allowed for specific comparisons to the age and sex specific population data from Norfolk, improving the accuracy of the results in chapter 4 and paper 5.3.

For the analyses in chapter 5, two important strengths were the flagging of all NOAR patients with the ONS and HSCIC, and the extensive follow up data within the cohort. These provide a sufficiently large number of mortality events to study different associations, as well as the rare opportunity to study mortality trends over time in RA. Few cohorts have sufficient follow up to undertake such mortality studies. Even in such a long running study, however, we were still limited to a 7 year period to capture mortality events in each cohort, as explained above.

A major weakness of a historical database such as NOAR is missing data. This may be due to loss to follow up, patients not attending follow up assessments, patients not completing questionnaires and even alterations in the variables collected over time. A variety of different statistical approaches were used to address the issue of missing data, depending on the individual analysis and as described in the methods and result chapters. However, it certainly remains possible that missing data have led to biased interpretation of the results. The influence of changes in data collection affected in particular the radiological data, where the time points when x-rays were taken changed between cohorts 1/2 and cohorts 3/4. These results should therefore be interpreted with some caution when looking at any changes over time. In addition, NOAR was not established with the particular aims and objectives of this thesis in mind. Therefore there may have been potentially important variables where either insufficient or no data were collected. For example, the analyses in chapter 6 would have been strengthened further if CRP had been measured more frequently to allow more frequent calculation of DAS28 scores. Equally, the availability of x-rays performed at regular intervals would have permitted better investigation of the association between radiological progression over time (as opposed to radiological damage alone) and satisfying the 2010 criteria or the presence or absence of autoantibodies.

A further weakness of NOAR in the context of thesis is the possibility of other diagnoses. At each follow up patient-reported diagnoses were collected, and patients were excluded after 5 years if an alternative diagnosis was made.

Consultant rheumatologist diagnosis was not otherwise routinely collected. In part this was due to the fact that in NOAR it has been noted that consultant diagnoses often changed over time and all patients were encompassed by the term IP. Nevertheless it is possible that some of the patients included had another inflammatory arthritis such as psoriatic arthritis, which could be a cause of poor outcome in seronegative patients or those who do not satisfy the 2010 criteria. Although formal classification criteria for PsA have not been applied in NOAR, psoriasis has previously been reported in approximately 9% of all patients with IP in NOAR, with similar rates in the RF negative subgroup (186). Patients with psoriasis also comprised only 9% of all patients who did not satisfy the 2010 criteria at baseline. Therefore PsA is unlikely to account for all patients who might be missed by the 2010 criteria but nevertheless have poor prognosis.

8.2.2 Collaborations

An important strength of the analyses in this thesis was in the use of validation cohorts and initiation of collaborative projects. Analyses in both chapters 5 and 6 were done in collaboration with Leiden University Medical Centre, Netherlands. The viewpoint article in chapter 7 was also the result of collaboration between all the co-authors. In chapter 5, this allowed the analyses to be replicated in an independent cohort to improve the generalisability of the results. This was important as there were associations which were seen in either NOAR or EAC separately but were not identified in the other cohort. The associations demonstrated in just one cohort may only be true for the specific group of patients within that cohort, and may not be relevant to the wider population of patients with RA and inflammatory arthritis. In chapter 6, the collaboration meant that NOAR blood samples could be tested for anti-CarP antibodies. In chapter 7, all co-authors shared their expert views and opinions which formed the basis of the manuscript.

8.2.3 Statistical methods

Assumptions were made in all of the various regression models employed in this thesis, as discussed in section 2.3 of the methods. They were selected as the most appropriate models at that time, given the individual objectives they were addressing and the nature and availability of the data within the cohort. In addition, throughout the analyses, careful efforts were made to account for any potential confounding. This included selecting more parsimonious models to avoid overadjustment bias, as described in section 2.2.3.

An important weakness of any observational cohort is the possibility of unmeasured confounding, which may be more likely in an historical cohort where the data has been collected prospectively but the analysis is done retrospectively. A possible example of an unmeasured confounder might be body composition, i.e. the distribution of fat and lean muscle mass within the body. In patients who are overweight or obese it can be more difficult to obtain an accurate swollen joint count; this could lead to them being potentially misclassified as having RA or not according to the 2010 criteria. If the outcome of interest is also associated with body composition, such as mortality, this may have been a confounder. Although proxy measures of body composition, such as body mass index (BMI) and waist circumference, were measured in NOAR, this was only done in a subgroup of patients, and no measurement of body composition was performed. It has been suggested that BMI is too crude a measure of body composition in patients with inflammatory arthritis (187), therefore, given there were also substantial missing data on this variable, it was not adjusted for in the analyses. However, it remains an example of the unmeasured confounding that may be present in the results presented here.

8.2.4 Erosive disease

It should be noted that since the publication of the 2010 criteria, a EULAR task force was established to define what is meant by 'typical RA erosion', and published their definition in 2013 (90). As this definition was not available at the outset of this thesis, and because few patients in NOAR have x-rays taken at baseline, the radiographic criterion was not applied in any of the analyses in this thesis. It should be noted that this is in keeping with the recommendations from the criteria publication, which state that it can be applied in patients where x-rays are already available, and they do not recommend that x-rays are taken to classify a patient (31). However it is possible that misclassification of patients may have occurred because of this approach. Notably, the effect of such misclassification will have been to underestimate the number of patients who satisfied the criteria.

8.3 Implications for clinical practice

The 2010 criteria were developed as a classification tool for research studies. As a result the analyses in this thesis which have contributed towards their validation have done so with that in mind. However there are clinical implications from a number of the results and the following recommendations are made:

- Although in general robust for classification for the purpose of research, the 2010 criteria should not be used as diagnostic criteria.
- The 2010 criteria may however aid diagnosis and be a useful guide in identifying patients with worse prognosis and those at increased risk of early mortality. These patients may benefit from more aggressive treatment strategies.
- Clinicians should be aware that patients who test negative for RF and ACPA may not be identified by the 2010 criteria and a subset of these may nevertheless have a poor prognosis.
- Patients with inflammatory arthritis who have more than one autoantibody may be at increased risk of early death. In addition, the measurement of anti-CarP antibodies is a promising test that can provide additional prognostic information to measurement of RF and ACPA. Therefore there may be clinical benefit in testing for more than one autoantibody to guide management decisions.

8.4 Future work

8.4.1 The 2010 ACR/EULAR classification criteria for RA in practice

A growing number of studies to date have employed the 2010 criteria in their primary role to define RA for research. The majority of these have been observational studies conducted retrospectively in cohorts with the available data such as the Leiden EAC (100) and the ESPOIR cohort (125), as well as the analyses in this thesis. They have begun to be incorporated as inclusion criteria for prospective cohort studies and RCTs (185) (138). Nevertheless they are not yet fully embedded in the rheumatology research psyche and some studies still select the 1987 criteria (188). Therefore the practical implications and usability of these criteria have not yet been tested on a wider scale. This will not require a specific

research analysis, but any challenges in routine use of the criteria will become apparent over time.

8.4.2 Validation in established disease

Formal, data-driven, validation of the 2010 criteria in patients with established inflammatory arthritis remains a notable gap in the literature. Application of the 2010 criteria in a prevalence study might be a useful and important first step towards this. It would give some idea of the practical implications of using the criteria in patients with longstanding disease, as well as updating the literature on an important measure of disease occurrence. It is however a significant undertaking, requiring a survey within a defined population (to provide a reference population for the denominator), and collection of data on all the parameters of the 2010 criteria in this population to ensure complete capture of cases.

8.4.3 Incorporation of MSK imaging

In the publication of the 2010 criteria the role of MSK imaging such as ultrasound and MRI was briefly mentioned, with a suggestion that they could be used to confirm clinical findings of the number of involved joints (31). A small number of studies, reviewed in the systematic review in section 1.8.8, have investigated whether imaging could be used to improve the performance of the criteria. It appears the most economical application of the imaging modalities is in patients who do not already satisfy the criteria by the clinical parameters. However, these results need confirming in larger studies and different populations. Further, although both MRI and US appear to be useful, it is not yet clear what the best setting or application is for each of these modalities.

8.4.4 Testing for multiple antibodies in inflammatory arthritis

Further work is also indicated to clarify the role of testing for more than one autoantibody in patients with inflammatory arthritis. The analyses in this thesis and others mentioned above (55;170;171) suggest there may be some merit in testing multiple antibodies, but which antibodies and how many may depend on the setting (clinical/research), the outcome of interest, as well as economic factors. In routine practice, patient often are tested for both RF and ACPA. However, how to interpret the prognostic implications of the results of both tests in combination is

not clear, and in the UK, some NHS trusts also place restrictions on the number of antibodies that can be tested. Currently anti-CarP antibodies can only be tested for in a research setting; however a commercial assay is being developed.

Overall it appears that despite more biomarkers being identified in inflammatory arthritis, and various classification criteria proposed which are useful for research purposes, practical stratification of patients for treatment remains challenging for clinicians. Currently, therefore, guidance recommends treating all patients aggressively from the outset, no matter what their antibody or classification criteria status is. Although there is no evidence to suggest this is harmful, it is possible that some patients are being exposed to treatments that they may not need. As the genetic studies have indicated that the presence or absence of different antibodies may represent different underlying pathogenesis (183), a possible strategy might be to analyse whether patients cluster based on their antibody profile as well as any other biomarkers that relate to underlying pathophysiological mechanisms. If successful, different treatment strategies could then be proposed and trialled based on these putative mechanisms, with the aim of developing a genuine stratified medicine treatment approach to inflammatory arthritis.

8.5 Final conclusions

In conclusion, this thesis has shown that the 2010 ACR/EULAR classification criteria have broadly good construct validity. They have been shown to identify more patients early in their disease course, and predict many of the adverse outcomes associated with poor prognosis in RA, such as mortality, disability, disease severity and radiological damage. The criteria may however miss a subgroup of seronegative patients with inflammatory arthritis at risk of poor outcomes. Testing for novel autoantibodies in these patients may prove useful.

Reference List

(1) Symmons D, Turner G, Webb R, Asten P, Barrett E, Lunt M, et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. Rheumatology (Oxford) 2002 Jul;41(7):793-800.

(2) Wolfe F. The natural history of rheumatoid arthritis. J Rheumatol Suppl 1996 Mar;44:13-22.

(3) Gordon P, West J, Jones H, Gibson T. A 10 year prospective followup of patients with rheumatoid arthritis 1986-96. J Rheumatol 2001 Nov;28(11):2409-15.

(4) Rasker JJ, Cosh JA. The natural history of rheumatoid arthritis over 20 years. Clinical symptoms, radiological signs, treatment, mortality and prognostic significance of early features. Clin Rheumatol 1987 Sep;6 Suppl 2:5-11.

(5) Duthie JJ, Brown PE, Truelove LH., Baragar FD, Lawrie AJ. Course and prognosis in rheumatoid arthritis. A further report. Ann Rheum Dis 1964 May;23:193-204.

(6) Symmons DP, Barrett EM, Bankhead CR, Scott DG, Silman AJ. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. Br J Rheumatol 1994 Aug;33(8):735-9.

(7) Avina-Zubieta JA, Thomas J, Sadatsafavi M, Lehman AJ, Lacaille D. Risk of incident cardiovascular events in patients with rheumatoid arthritis: a meta-analysis of observational studies. Ann Rheum Dis 2012 Sep;71(9):1524-9.

(8) Mercer LK, Davies R, Galloway JB, Low A, Lunt M, Dixon WG, et al. Risk of cancer in patients receiving non-biologic disease-modifying therapy for rheumatoid arthritis compared with the UK general population. Rheumatology (Oxford) 2013 Jan;52(1):91-8.

(9) Prete M, Racanelli V, Digiglio L, Vacca A, Dammacco F, Perosa F. Extra-articular manifestations of rheumatoid arthritis: An update. Autoimmun Rev 2011 Dec;11(2):123-31.

(10) van der Woude D, Houwing-Duistermaat JJ, Toes RE, Huizinga TW, Thomson W, Worthington J, et al. Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. Arthritis Rheum 2009 Apr;60(4):916-23.

(11) Pedersen M, Jacobsen S, Garred P, Madsen HO, Klarlund M, Svejgaard A, et al. Strong combined gene-environment effects in anti-cyclic citrullinated peptide-positive rheumatoid arthritis: a nationwide case-control study in Denmark. Arthritis Rheum 2007 May;56(5):1446-53.

(12) Lahiri M, Luben RN, Morgan C, Bunn DK, Marshall T, Lunt 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 Jan;73(1):219-26.

(13) Harrison BJ, Silman AJ, Barrett EM, Scott DG, Symmons DP. Presence of psoriasis does not influence the presentation or short-term outcome of patients with early inflammatory polyarthritis. J Rheumatol 1997 Sep;24(9):1744-9.

(14) Pedersen M, Jacobsen S, Klarlund M, Pedersen BV, Wiik A, Wohlfahrt J, et al. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. Arthritis Res Ther 2006;8(4):R133.

(15) Karlson EW, Mandl LA, Hankinson SE, Grodstein F. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? Results from the Nurses' Health Study. Arthritis Rheum 2004 Nov;50(11):3458-67.

(16) Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. Lancet 2010 Sep 25;376(9746):1094-108.

(17) Verpoort KN, Van Dongen H, Allaart CF, Toes RE, Breedveld FC, Huizinga TW. Undifferentiated arthritis--disease course assessed in several inception cohorts. Clin Exp Rheumatol 2004 Sep;22(5 Suppl 35):S12-S17.

(18) Van Aken J, Van Dongen H, le Cessie S, Allaart CF, Breedveld FC, Huizinga TW. 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 Jan;65(1):20-5.

(19) Berthelot JM, Klarlund M, McGonagle D, Bernelot-Moens HJ, Calin A, Harrison B, et al. Lessons from an international survey of paper cases of 10 real patients from an early arthritis clinic. CRI (Club Rhumatismes et Inflammation) Group. J Rheumatol 2001 May;28(5):975-81.

(20) Harrison BJ, Symmons DP, Brennan P, Barrett EM, Silman AJ. Natural remission in inflammatory polyarthritis: issues of definition and prediction. Br J Rheumatol 1996 Nov;35(11):1096-100.

(21) Tunn EJ, Bacon PA. Differentiating persistent from self-limiting symmetrical synovitis in an early arthritis clinic. Br J Rheumatol 1993 Feb;32(2):97-103.

(22) Wiles N, Symmons DP, Harrison B, Barrett E, Barrett JH, Scott DG, et al. Estimating the incidence of rheumatoid arthritis: trying to hit a moving target? Arthritis Rheum 1999 Jul;42(7):1339-46.

(23) del Puente A, Knowler WC, Pettitt DJ, Bennett PH. The incidence of rheumatoid arthritis is predicted by rheumatoid factor titer in a longitudinal population study. Arthritis Rheum 1988 Oct;31(10):1239-44.

(24) Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003 Oct;48(10):2741-9.

(25) Bos WH, Wolbink GJ, Boers M, Tijhuis GJ, de Vries N, van der Horst-Bruinsma IE, et al. Arthritis development in patients with arthralgia is strongly associated with anti-citrullinated protein antibody status: a prospective cohort study. Ann Rheum Dis 2010 Mar;69(3):490-4.

(26) Brink M, Hansson M, Mathsson L, Jakobsson PJ, Holmdahl R, Hallmans G, et al. Multiplex analyses of antibodies against citrullinated peptides in individuals prior to development of rheumatoid arthritis. Arthritis Rheum 2013 Apr;65(4):899-910.

(27) Dorner T, Egerer K, Feist E, Burmester GR. Rheumatoid factor revisited. Curr Opin Rheumatol 2004 May;16(3):246-53.

(28) Waaler E. On the occurrence of a factor in human serum activating the specific agglutination of sheep blood corpuscles. Acta Pathologica Microbiologica Scandinavica 1940 May 1;17(2):172-88.

(29) Ropes MW, Bennett GA, Cobb S, Jacox R, Jessar RA. 1958 Revision of diagnostic criteria for rheumatoid arthritis. Bull Rheum Dis 1958 Dec;9(4):175-6.

(30) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988 Mar;31(3):315-24.

(31) Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, III, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010 Sep;62(9):2569-81.

(32) Visser H, Gelinck LB, Kampfraath AH, Breedveld FC, Hazes JM. Diagnostic and prognostic characteristics of the enzyme linked immunosorbent rheumatoid factor assays in rheumatoid arthritis. Ann Rheum Dis 1996 Mar;55(3):157-61.

(33) Teitsson I, Withrington RH, Seifert MH, Valdimarsson H. Prospective study of early rheumatoid arthritis. I. Prognostic value of IgA rheumatoid factor. Ann Rheum Dis 1984 Oct;43(5):673-8.

(34) Young A, Koduri G. Extra-articular manifestations and complications of rheumatoid arthritis. Best Pract Res Clin Rheumatol 2007 Oct;21(5):907-27.

(35) Turesson C, O'Fallon WM, Crowson CS, Gabriel SE, Matteson EL. Extra-articular disease manifestations in rheumatoid arthritis: incidence trends and risk factors over 46 years. Ann Rheum Dis 2003 Aug;62(8):722-7.

(36) Barra L, Bykerk V, Pope JE, Haraoui BP, Hitchon CA, Thorne JC, et al. Anticitrullinated protein antibodies and rheumatoid factor fluctuate in early inflammatory arthritis and do not predict clinical outcomes. J Rheumatol 2013 Aug;40(8):1259-67.

(37) Goodson NJ, Wiles NJ, Lunt M, Barrett EM, Silman AJ, Symmons DP. Mortality in early inflammatory polyarthritis: cardiovascular mortality is increased in seropositive patients. Arthritis Rheum 2002 Aug;46(8):2010-9.

(38) Gonzalez A, Icen M, Kremers HM, Crowson CS, Davis JM, III, Therneau TM, et al. Mortality trends in rheumatoid arthritis: the role of rheumatoid factor. J Rheumatol 2008 Jun;35(6):1009-14.

(39) Ingegnoli F, Castelli R, Gualtierotti R. Rheumatoid factors: clinical applications. Dis Markers 2013;35(6):727-34.

(40) Simard JF, Holmqvist M. Rheumatoid factor positivity in the general population. BMJ 2012;345:e5841.

(41) Dequeker J, Van NR, Vandepitte J. Age-related rheumatoid factors. Incidence and characteristics. Ann Rheum Dis 1969 Jul;28(4):431-6.

(42) Ursum J, Bos WH, van DN, Dijkmans BA, Van SD. Levels of anti-citrullinated protein antibodies and IgM rheumatoid factor are not associated with outcome in early arthritis patients: a cohort study. Arthritis research & therapy 2010;12(1):R8.

(43) Farragher TM, Lunt M, Fu B, Bunn D, Symmons DPM. Early treatment with, and time receiving, first disease-modifying antirheumatic drug predicts long-term function in patients with inflammatory polyarthritis. Ann Rheum Dis 2010;69(4):689-95.

(44) Nell VP, Machold KP, Stamm TA, Eberl G, Heinzl H, Uffmann M, et al. Autoantibody profiling as early diagnostic and prognostic tool for rheumatoid arthritis. Ann Rheum Dis 2005 Dec;64(12):1731-6.

(45) Aletaha D, Alasti F, Smolen JS. Rheumatoid factor determines structural progression of rheumatoid arthritis dependent and independent of disease activity. Ann Rheum Dis 2013 Jun;72(6):875-80.

(46) van Venrooij WJ, van Beers JJ, Pruijn GJ. Anti-CCP antibodies: the past, the present and the future. Nat Rev Rheumatol 2011 Jul;7(7):391-8.

(47) Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 2006 Jan;54(1):38-46.

(48) Nienhuis RL, Mandema E. A new serum factor in patients with rheumatoid arthritis; the antiperinuclear factor. Ann Rheum Dis 1964 Jul;23:302-5.

(49) Sebbag M, Simon M, Vincent C, Masson-Bessiere C, Girbal E, Durieux JJ, et al. The antiperinuclear factor and the so-called antikeratin antibodies are the same rheumatoid arthritis-specific autoantibodies. J Clin Invest 1995 Jun;95(6):2672-9.

(50) Schellekens GA, Visser H, de Jong BA, van den Hoogen FH, Hazes JM, Breedveld FC, et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum 2000 Jan;43(1):155-63.

(51) Harre U, Georgess D, Bang H, Bozec A, Axmann R, Ossipova E, et al. Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin. J Clin Invest 2012 May;122(5):1791-802.

(52) Kastbom A, Strandberg G, Lindroos A, Skogh T. Anti-CCP antibody test predicts the disease course during 3 years in early rheumatoid arthritis (the Swedish TIRA project). Ann Rheum Dis 2004 Sep;63(9):1085-9.

(53) Farragher TM, Goodson NJ, Naseem H, Silman AJ, Thomson W, Symmons D, et al. Association of the HLA-DRB1 gene with premature death, particularly from cardiovascular disease, in patients with rheumatoid arthritis and inflammatory polyarthritis. Arthritis Rheum 2008 Feb;58(2):359-69.

(54) Machold KP, Stamm TA, Nell VP, Pflugbeil S, Aletaha D, Steiner G, et al. Very recent onset rheumatoid arthritis: clinical and serological patient characteristics associated with radiographic progression over the first years of disease. Rheumatology (Oxford) 2007 Feb;46(2):342-9.

(55) van der Linden MP, van der Woude D, Ioan-Facsinay A, Levarht EW, Stoeken-Rijsbergen G, Huizinga TW, et al. Value of anti-modified citrullinated vimentin and third-generation anti-cyclic citrullinated peptide compared with second-generation anti-cyclic citrullinated peptide and rheumatoid factor in predicting disease outcome in undifferentiated arthritis and rheumatoid arthritis. Arthritis Rheum 2009 Aug;60(8):2232-41.

(56) van den Broek M, Dirven L, Klarenbeek NB, Molenaar TH, Han KH, Kerstens PJ, et al. The association of treatment response and joint damage with ACPA-status in recent-onset RA: a subanalysis of the 8-year follow-up of the BeSt study. Ann Rheum Dis 2012 Feb;71(2):245-8.

(57) Farragher TM, Lunt M, Plant D, Bunn DK, Barton A, Symmons DP. Benefit of early treatment in inflammatory polyarthritis patients with anti-cyclic citrullinated peptide antibodies versus those without antibodies. Arthritis care & research 2010 May;62(5):664-75.

(58) Shi J, van de Stadt LA, Levarht EW, Huizinga TW, Toes RE, Trouw LA, et al. Anti-carbamylated protein antibodies are present in arthralgia patients and

predict the development of rheumatoid arthritis. Arthritis Rheum 2013 Apr;65(4):911-5.

(59) Shi J, van de Stadt LA, Levarht EW, Huizinga TW, Hamann D, Van SD, et al. Anti-carbamylated protein (anti-CarP) antibodies precede the onset of rheumatoid arthritis. Ann Rheum Dis 2014 Apr;73(4):780-3.

(60) Trouw LA, Huizinga TW, Toes RE. Autoimmunity in rheumatoid arthritis: different antigens--common principles. Ann Rheum Dis 2013 Apr;72 Suppl 2:ii132-ii136.

(61) Shi J, Knevel R, Suwannalai P, van der Linden MP, Janssen GM, van Veelen PA, et al. Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. Proc Natl Acad Sci U S A 2011 Oct 18;108(42):17372-7.

(62) Wilske KR, Healey LA. Remodeling the pyramid--a concept whose time has come. J Rheumatol 1989 May;16(5):565-7.

(63) Constable TJ, Crockson RA, Crockson AP, McConkey B. Drug treatment of rheumatoid arthritis. A systematic approach. Lancet 1975 May 24;1(7917):1176-80.

(64) Wilke WS, Sweeney TJ, Calabrese LH. Early, aggressive therapy for rheumatoid arthritis: concerns, descriptions, and estimate of outcome. Semin Arthritis Rheum 1993 Oct;23(2 Suppl 1):26-41.

(65) Ward JR. Role of disease-modifying antirheumatic drugs versus cytotoxic agents in the therapy of rheumatoid arthritis. Am J Med 1988 Oct 14;85(4A):39-44.

(66) Furst DE, Kremer JM. Methotrexate in rheumatoid arthritis. Arthritis Rheum 1988 Mar;31(3):305-14.

(67) Verstappen SM, Jacobs JW, Bijlsma JW, Heurkens AH, van Booma-Frankfort C, Borg EJ, et al. Five-year followup of rheumatoid arthritis patients after early treatment with disease-modifying antirheumatic drugs versus treatment according to the pyramid approach in the first year. Arthritis Rheum 2003 Jul;48(7):1797-807.

(68) Boers M, Verhoeven AC, Markusse HM, van de Laar MA, Westhovens R, van Denderen JC, et al. Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. Lancet 1997 Aug 2;350(9074):309-18.

(69) Quinn MA, Emery P. Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. Clin Exp Rheumatol 2003 Sep;21(5 Suppl 31):S154-S157.

(70) Nell VP, Machold KP, Eberl G, Stamm TA, Uffmann M, Smolen JS. Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis. Rheumatology (Oxford) 2004 Jul;43(7):906-14.

(71) Rau R, Schleusser B, Herborn G, Karger T. Long-term treatment of destructive rheumatoid arthritis with methotrexate. J Rheumatol 1997 Oct;24(10):1881-9.

(72) Wluka A, Buchbinder R, Mylvaganam A, Hall S, Harkness A, Lewis D, et al. Longterm methotrexate use in rheumatoid arthritis: 12 year followup of 460 patients treated in community practice. J Rheumatol 2000 Aug;27(8):1864-71.

(73) Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, Van Zeben D, Kerstens PJ, Hazes JM, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. Arthritis Rheum 2005 Nov;52(11):3381-90.

(74) Landewe RB, Boers M, Verhoeven AC, Westhovens R, van de Laar MA, Markusse HM, et al. COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention. Arthritis Rheum 2002 Feb;46(2):347-56.

(75) Nam JL, Ramiro S, Gaujoux-Viala C, Takase K, Leon-Garcia M, Emery P, et al. Efficacy of biological disease-modifying antirheumatic drugs: a systematic literature review informing the 2013 update of the EULAR recommendations for the management of rheumatoid arthritis. Ann Rheum Dis 2014 Mar;73(3):516-28.

(76) Harrison BJ, Symmons DP, Barrett EM, Silman AJ. The performance of the 1987 ARA classification criteria for rheumatoid arthritis in a population based cohort of patients with early inflammatory polyarthritis. American Rheumatism Association. J Rheumatol 1998 Dec;25(12):2324-30.

(77) Ropes MW, Bennett GA, Cobb S, Jacox R, Jessar RA. Proposed diagnostic criteria for rheumatoid arthritis; report of a study conducted by a committee of the American Rheumatism Association. J Chronic Dis 1957 Jun;5(6):630-5.

(78) Cobb S, Thompson DJ, Rosenbaum J, Warren JE, Merchant WR. On the measurement of prevalence of arthritis and rheumatism from interview data. J Chronic Dis 1956 Feb;3(2):134-9.

(79) Pincus T, Callahan LF. Reassessment of twelve traditional paradigms concerning the diagnosis, prevalence, morbidity and mortality of rheumatoid arthritis. Scand J Rheumatol Suppl 1989;79:67-96.

(80) Bernelot Moens HJ, van de Laar MA, van der Korst JK. Comparison of the sensitivity and specificity of the 1958 and 1987 criteria for rheumatoid arthritis. J Rheumatol 1992 Feb;19(2):198-203.

(81) Saraux A, Berthelot JM, Chales G, Le Henaff C, Thorel JB, Hoang S, et al. Ability of the American College of Rheumatology 1987 criteria to predict rheumatoid arthritis in patients with early arthritis and classification of these patients two years later. Arthritis Rheum 2001 Nov;44(11):2485-91.

(82) Van Dongen H, Van Aken J, Lard LR, Visser K, Ronday HK, Hulsmans HM, et al. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. Arthritis Rheum 2007 May;56(5):1424-32.

(83) Saunders SA, Capell HA, Stirling A, Vallance R, Kincaid W, McMahon AD, et al. Triple therapy in early active rheumatoid arthritis: a randomized, single-blind, controlled trial comparing step-up and parallel treatment strategies. Arthritis Rheum 2008 May;58(5):1310-7.

(84) Rantalaiho V, Korpela M, Laasonen L, Kautiainen H, Jarvenpaa S, Hannonen P, et al. Early combination disease-modifying antirheumatic drug therapy and tight disease control improve long-term radiologic outcome in patients with early rheumatoid arthritis: the 11-year results of the Finnish Rheumatoid Arthritis Combination Therapy trial. Arthritis Res Ther 2010;12(3):R122.

(85) Funovits J, Aletaha D, Bykerk V, Combe B, Dougados M, Emery P, et al. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: methodological report phase I. Ann Rheum Dis 2010 Sep;69(9):1589-95.

(86) Neogi T, Aletaha D, Silman AJ, Naden RL, Felson DT, Aggarwal R, et al. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: Phase 2 methodological report. Arthritis Rheum 2010 Sep;62(9):2582-91.

(87) Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. Semin Arthritis Rheum 2011 Oct;41(2):95-105.

(88) Black N, Murphy M, Lamping D, McKee M, Sanderson C, Askham J, et al. Consensus development methods: a review of best practice in creating clinical guidelines. J Health Serv Res Policy 1999 Oct;4(4):236-48.

(89) Knevel R, Lukas C, van der Heijde D, Rincheval N, Combe B, van der Helm-van Mil AH. Defining erosive disease typical of RA in the light of the ACR/EULAR 2010 criteria for rheumatoid arthritis; results of the data driven phase. Ann Rheum Dis 2013 Apr;72(4):590-5.

(90) van der Heijde D, van der Helm-van Mil AH, Aletaha D, Bingham CO, Burmester GR, Dougados M, et al. EULAR definition of erosive disease in light of the 2010 ACR/EULAR rheumatoid arthritis classification criteria. Ann Rheum Dis 2013 Apr;72(4):479-81.

(91) Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000 Feb 29;19(4):453-73.

(92) Alves C, Luime JJ, Van Zeben D, Huisman AM, Weel AE, Barendregt PJ, et al. Diagnostic performance of the ACR/EULAR 2010 criteria for rheumatoid arthritis and two diagnostic algorithms in an early arthritis clinic (REACH). Ann Rheum Dis 2011 Sep;70(9):1645-7.

(93) Britsemmer K, Ursum J, Gerritsen M, van Tuyl LH, Van SD. Validation of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: slight improvement over the 1987 ACR criteria. Ann Rheum Dis 2011 Aug;70(8):1468-70.

(94) Cader MZ, Filer A, Hazlehurst J, de PP, Buckley CD, Raza K. Performance of the 2010 ACR/EULAR criteria for rheumatoid arthritis: comparison with 1987 ACR criteria in a very early synovitis cohort. Ann Rheum Dis 2011 Jun;70(6):949-55.

(95) Cornec D, Varache S, Morvan J, vauchelle-Pensec Vr, Berthelot JM, Le Henaff-Bourhis C, et al. Comparison of ACR 1987 and ACR/EULAR 2010 criteria for predicting a 10-year diagnosis of rheumatoid arthritis. Joint Bone Spine 2012 Dec;79(6):581-5.

(96) Kaarela K, Kauppi JE, Kauppi MJ. The 2010 ACR/EULAR classification criteria for rheumatoid arthritis in the Heinola inception cohort--diagnoses confirmed by long-term follow-up. Clin Rheumatol 2012 Mar;31(3):547-51.

(97) Kaneko Y, Kuwana M, Kameda H, Takeuchi T. Sensitivity and specificity of 2010 rheumatoid arthritis classification criteria. Rheumatology (Oxford) 2011 Jul;50(7):1268-74.

(98) van de Sande MG, de Hair MJ, Schuller Y, van de Sande GP, Wijbrandts CA, Dinant HJ, et al. The features of the synovium in early rheumatoid arthritis according to the 2010 ACR/EULAR classification criteria. PLoS One 2012;7(5):e36668.

(99) Mourao AF, Canhao H, Moura RA, Cascao R, Weinmann P, Rodrigues A, et al. Markers of progression to rheumatoid arthritis: discriminative value of the new ACR/EULAR rheumatoid arthritis criteria in a portuguese population with early polyarthritis. Acta Reumatol Port 2011 Oct;36(4):370-6.

(100) van der Linden MP, Knevel R, Huizinga TW, van der Helm-van Mil A. Classification of rheumatoid arthritis: comparison of the 1987 American College of Rheumatology criteria and the 2010 American College of Rheumatology/European League Against Rheumatism criteria. Arthritis Rheum 2011 Jan;63(1):37-42.

(101) Varache S, Cornec D, Morvan J, vauchelle-Pensec V, Berthelot JM, Le Henaff-Bourhis C, et al. Diagnostic accuracy of ACR/EULAR 2010 criteria for rheumatoid arthritis in a 2-year cohort. J Rheumatol 2011 Jul;38(7):1250-7.

(102) van der Helm-van Mil AHM, Huizinga TWH. The Key to Early Rheumatoid Arthritis. The Rheumatologist [September 2010]. 2010.

Ref Type: Magazine Article

(103) Sokka T, Abelson B, Pincus T. Mortality in rheumatoid arthritis: 2008 update. Clin Exp Rheumatol 2008 Sep;26(5 Suppl 51):S35-S61.

(104) Smolina K, Wright FL, Rayner M, Goldacre MJ. Determinants of the decline in mortality from acute myocardial infarction in England between 2002 and 2010: linked national database study. BMJ 2012;344:d8059.

(105) Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis 2014 Mar;73(3):492-509.

(106) Combe B, Rincheval N, Benessiano J, Berenbaum F, Cantagrel A, Daures JP, et al. Five-year favorable outcome of patients with early rheumatoid arthritis in the 2000s: data from the ESPOIR cohort. J Rheumatol 2013 Oct;40(10):1650-7.

(107) Berglin E, Dahlqvist SR. Comparison of the 1987 ACR and 2010 ACR/EULAR classification criteria for rheumatoid arthritis in clinical practice: a prospective cohort study. Scand J Rheumatol 2013;42(5):362-8.

(108) Biliavska I, Stamm TA, Martinez-Avila J, Huizinga TW, Landewe RB, Steiner G, et al. Application of the 2010 ACR/EULAR classification criteria in patients with very early inflammatory arthritis: analysis of sensitivity, specificity and predictive values in the SAVE study cohort. Ann Rheum Dis 2012 Sep 15.

(109) De Hair MJH, Lehmann KA, Van De Sande MGH, Maijer KI, Gerlag DM, Tak PP. The clinical picture of rheumatoid arthritis according to the 2010 American College of Rheumatology/European League Against Rheumatism criteria: Is this still the same disease? Arthritis Rheum64(2):February.

(110) Kasturi S, Goldstein BL, Malspeis S, Karlson EW, Costenbader KH. Comparison of the 1987 American College of Rheumatology and the 2010 American College of Rheumatology/European League against Rheumatism criteria for classification of rheumatoid arthritis in the Nurses' Health Study cohorts. Rheumatol Int 2014 Mar;34(3):407-11.

(111) Kawashiri Sy, Suzuki T, Okada A, Yamasaki S, Tamai M, Nakamura H, et al. Musculoskeletal ultrasonography assists the diagnostic performance of the 2010 classification criteria for rheumatoid arthritis. Mod Rheumatol 2012;1-8.

(112) Kennish L, Labitigan M, Budoff S, Filopoulos MT, McCracken WA, Swearingen CJ, et al. Utility of the new rheumatoid arthritis 2010 ACR/EULAR classification criteria in routine clinical care. BMJ Open 2012 Oct;2(5).

(113) Kim JY, Cho SK, Han M, Choi YY, Bae SC, Sung YK. The role of bone scintigraphy in the diagnosis of rheumatoid arthritis according to the 2010 ACR/EULAR classification criteria. J Korean Med Sci 2014 Feb;29(2):204-9.

(114) Makinen H, Kaarela K, Huhtala H, Hannonen PJ, Korpela M, Sokka T. Do the 2010 ACR/EULAR or ACR 1987 classification criteria predict erosive disease in early arthritis? Ann Rheum Dis 2012 Nov 21.

(115) Nakagomi D, Ikeda K, Okubo A, Iwamoto T, Sanayama Y, Takahashi K, et al. Ultrasound can improve the accuracy of the 2010 American College of Rheumatology/European League against rheumatism classification criteria for rheumatoid arthritis to predict the requirement for methotrexate treatment. Arthritis Rheum 2013 Apr;65(4):890-8.

(116) Nieuwenhuis WP, Reijnierse M, van der Helm-van Mil AH. Does adding the presence of MRI detected bone marrow oedema improve the accuracy of the 2010 EULAR/ACR criteria for rheumatoid arthritis? Ann Rheum Dis 2015 Mar;74(3):e29.

(117) Raja R, Chapman PT, O'Donnell JL, Ipenburg J, Frampton C, Hurst M, et al. Comparison of the 2010 american college of rheumatology/european league against rheumatism and the 1987 american rheumatism association classification criteria for rheumatoid arthritis in an early arthritis cohort in new zealand. J Rheumatol 2012 Nov;39(11):2098-103.

(118) Ravindran V, Abdulaziz A, Bhargavan PV. Performance of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis in a prospective early arthritis cohort in Kerala, India. Indian Journal of Rheumatology 2014 Sep;9(3):107-11.

(119) Reneses S, Pestana L, Garcia A. Comparison of the 1987 ACR criteria and the 2010 ACR/EULAR criteria in an inception cohort of patients with recent-onset inflammatory polyarthritis. Clin Exp Rheumatol 2012 May;30(3):417-20.

(120) Tamai M, Kita J, Nakashima Y, Suzuki T, Horai Y, Okada A, et al. Combination of MRI-detected bone marrow oedema with 2010 rheumatoid arthritis classification criteria improves the diagnostic probability of early rheumatoid arthritis. Ann Rheum Dis 2014 Dec;73(12):2219-20.

(121) Tamas MM, Felea I, Rednic S. How much difference does the age at onset make in early arthritis patients? Comparison between the ACR 1987 and the ACR/EULAR 2010 classification criteria for rheumatoid arthritis at the time of diagnosis. Rheumatol Int 2012 Sep 7.

(122) Zhao J, Li G, Xu J, Zhang X, Zhang Z. Performance of the American College of Rheumatology/European League Against Rheumatism 2010 criteria for the diagnosis of rheumatoid arthritis in Chinese patients. Joint, bone, spine : revue du rhumatisme 2013;80(5):482-7.

(123) Burgers LE, van Nies JA, Ho LY, de Rooy DP, Huizinga TW, van der Helm-van Mil AH. Long-term outcome of rheumatoid arthritis defined according to the 2010-classification criteria. Ann Rheum Dis 2014 Feb;73(2):428-32.

(124) Bykerk VP, Jamal S, Boire G, Hitchon CA, Haraoui B, Pope JE, et al. The Canadian Early Arthritis Cohort (CATCH): Patients with New-onset Synovitis Meeting the 2010 ACR/EULAR Classification Criteria But Not the 1987 ACR Classification Criteria Present with Less Severe Disease Activity. J Rheumatol 2012 Nov;39(11):2071-80.

(125) Fautrel B, Combe B, Rincheval N, Dougados M. Level of agreement of the 1987 ACR and 2010 ACR/EULAR rheumatoid arthritis classification criteria: an analysis based on ESPOIR cohort data. Ann Rheum Dis 2012 Mar;71(3):386-9.

(126) Krabben A, Huizinga TW, van der Helm-van Mil AH. Undifferentiated arthritis characteristics and outcomes when applying the 2010 and 1987 criteria for rheumatoid arthritis. Ann Rheum Dis 2012 Feb;71(2):238-41.

(127) Jung SJ, Kang Y, Ha YJ, Lee KH, Lee SW, Lee SK, et al. Application of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis in Korean patients with undifferentiated arthritis. Scand J Rheumatol 2012 May;41(3):192-5.

(128) Mueller RB, Schiff M, Kaegi T, Finckh A, Haile SR, Schulze-Koops H, et al. The new 2010 ACR/EULAR criteria as predictor of clinical and radiographic response in patients with early arthritis. Clin Rheumatol 2015 Jan;34(1):51-9.

(129) Saraux A, Tobon GJ, Benhamou M, Devauchelle-Pensec V, Dougados M, Mariette X, et al. Potential classification criteria for rheumatoid arthritis after two years: results from a French multicenter cohort. Arthritis Care Res (Hoboken) 2013 Aug;65(8):1227-34.

(130) Tamas MM, Rednic N, Felea I, Rednic S. Ultrasound assessment for the rapid classification of early arthritis patients. J Investig Med 2013 Dec;61(8):1184-91.

(131) van der Linden MP, Batstra MR, Bakker-Jonges LE, Detert J, Bastian H, Scherer HU, et al. Toward a data-driven evaluation of the 2010 American College of Rheumatology/European League Against Rheumatism criteria for rheumatoid arthritis: is it sensible to look at levels of rheumatoid factor? Arthritis Rheum 2011 May;63(5):1190-9.

(132) Galloway JB, Hyrich KL, Mercer LK, Dixon WG, Ustianowski AP, Helbert M, et al. Risk of septic arthritis in patients with rheumatoid arthritis and the effect of anti-TNF therapy: results from the British Society for Rheumatology Biologics Register. Ann Rheum Dis 2011 Oct;70(10):1810-4.

(133) Ramiro S, Gaujoux-Viala C, Nam JL, Smolen JS, Buch M, Gossec L, et al. Safety of synthetic and biological DMARDs: a systematic literature review informing the 2013 update of the EULAR recommendations for management of rheumatoid arthritis. Ann Rheum Dis 2014 Mar;73(3):529-35.

(134) de Launay D, van de Sande MG, de Hair MJ, Grabiec AM, van de Sande GP, Lehmann KA, et al. Selective involvement of ERK and JNK mitogen-activated protein kinases in early rheumatoid arthritis (1987 ACR criteria compared to 2010 ACR/EULAR criteria): a prospective study aimed at identification of diagnostic and prognostic biomarkers as well as therapeutic targets. Ann Rheum Dis 2012 Mar 1;71(3):415-23.

(135) Jung SJ, Lee SW, Ha YJ, Lee KH, Kang Y, Park MC, et al. Patients with early arthritis who fulfil the 1987 ACR classification criteria for rheumatoid arthritis but not the 2010 ACR/EULAR criteria. Ann Rheum Dis 2012 Jun;71(6):1097-8.

(136) van der Helm-van Mil A, le Cessie S, Van Dongen H, Breedveld FC, Toes RE, Huizinga TW. A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. Arthritis Rheum 2007 Feb;56(2):433-40.

(137) Yazici H. Diagnostic versus classification criteria - a continuum. Bull NYU Hosp Jt Dis 2009;67(2):206-8.

(138) Haschka J, Englbrecht M, Hueber AJ, Manger B, Kleyer A, Reiser M, et al. Relapse rates in patients with rheumatoid arthritis in stable remission tapering or stopping antirheumatic therapy: interim results from the prospective randomised controlled RETRO study. Ann Rheum Dis 2015 Feb 6.

(139) Symmons DP, Silman AJ. The Norfolk Arthritis Register (NOAR). Clin Exp Rheumatol 2003 Sep;21(5 Suppl 31):S94-S99.

(140) Symmons DP, Silman AJ. Aspects of early arthritis. What determines the evolution of early undifferentiated arthritis and rheumatoid arthritis? An update from the Norfolk Arthritis Register. Arthritis Res Ther 2006;8(4):214.

(141) Wolfe F, Kleinheksel SM, Cathey MA, Hawley DJ, Spitz PW, Fries JF. The clinical value of the Stanford Health Assessment Questionnaire Functional Disability Index in patients with rheumatoid arthritis. J Rheumatol 1988 Oct;15(10):1480-8.

(142) Larson A, Dale K, Eek M. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. Acta Radiol Diagn (Stockh) 1977;18:481-91.

(143) World Health Organization. International statistical classification of diseases and related health problems 10th revision (ICD-10). : 2010Available from: URL: http://apps.who.int/classifications/icd10/browse/2010/en

(144) Mirjafari H. The prevalence and determinants of subclinical atherosclerosis in an early inflammatory polyarthritis inception cohort. University of Manchester; 2011.

(145) Uhlig T, Smedstad LM, Vaglum P, Moum T, Gerard N, Kvien TK. The course of rheumatoid arthritis and predictors of psychological, physical and radiographic outcome after 5 years of follow-up. Rheumatology (Oxford) 2000;39(7):732-41.

(146) Mjaavatten MD, van der Heijde DM, Uhlig T, Haugen AJ, Nygaard H, Bjorneboe O, et al. Should anti-citrullinated protein antibody and rheumatoid factor status be reassessed during the first year of followup in recent-onset arthritis? A longitudinal study. J Rheumatol 2011 Nov;38(11):2336-41.

(147) Barra L, Pope J, Bessette L, Haraoui B, Bykerk V. Lack of seroconversion of rheumatoid factor and anti-cyclic citrullinated peptide in patients with early inflammatory arthritis: a systematic literature review. Rheumatology (Oxford) 2011 Feb;50(2):311-6.

(148) Arend WP, Firestein GS. Pre-rheumatoid arthritis: predisposition and transition to clinical synovitis. Nat Rev Rheumatol 2012 Oct;8(10):573-86.

(149) Robins J. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. J Chronic Dis 1987;40 Suppl 2:139S-61S.

(150) Robins JM. Data, design, and background knowledge in etiologic inference. Epidemiology 2001 May;12(3):313-20.

(151) Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology 2009 Jul;20(4):488-95.

(152) Sugiyama D, Nishimura K, Tamaki K, Tsuji G, Nakazawa T, Morinobu A, et al. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. Ann Rheum Dis 2010 Jan;69(1):70-81.

(153) Groenwold RH, Nelson DB, Nichol KL, Hoes AW, Hak E. Sensitivity analyses to estimate the potential impact of unmeasured confounding in causal research. Int J Epidemiol 2010 Feb;39(1):107-17.

(154) Glynn RJ, Schneeweiss S, Sturmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. Basic Clin Pharmacol Toxicol 2006 Mar;98(3):253-9.

(155) Little RJA, Rubin DB. Statistical analysis with missing data. New York: Wiley; 1987.

(156) White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011 Feb 20;30(4):377-99.

(157) Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 1986 Mar;42(1):121-30.

(158) Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982 Dec;38(4):963-74.

(159) Gardiner JC, Luo Z, Roman LA. Fixed effects, random effects and GEE: what are the differences? Stat Med 2009 Jan 30;28(2):221-39.

(160) Diffin JG, Lunt M, Marshall T, Chipping JR, Symmons DP, Verstappen SM. Has the severity of rheumatoid arthritis at presentation diminished over time? J Rheumatol 2014 Aug;41(8):1590-9.

(161) Eriksson JK, Neovius M, Ernestam S, Lindblad S, Simard JF, Askling J. Incidence of rheumatoid arthritis in Sweden: a nationwide population-based assessment of incidence, its determinants, and treatment penetration. Arthritis Care Res (Hoboken) 2013 Jun;65(6):870-8.

(162) Rossini M, Rossi E, Bernardi D, Viapiana O, Gatti D, Idolazzi L, et al. Prevalence and incidence of rheumatoid arthritis in Italy. Rheumatol Int 2014 May;34(5):659-64.

(163) Kuo CF, Luo SF, See LC, Chou IJ, Chang HC, Yu KH. Rheumatoid arthritis prevalence, incidence, and mortality rates: a nationwide population study in Taiwan. Rheumatol Int 2013 Feb;33(2):355-60.

(164) Zeng QY, Chen R, Darmawan J, Xiao ZY, Chen SB, Wigley R, et al. Rheumatic diseases in China. Arthritis Res Ther 2008;10(1):R17.

(165) Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. Arthritis Res 2002;4 Suppl 3:S265-S272.

(166) Lau E, Symmons D, Bankhead C, MacGregor A, Donnan S, Silman A. Low prevalence of rheumatoid arthritis in the urbanized Chinese of Hong Kong. J Rheumatol 1993 Jul;20(7):1133-7.

(167) Doran MF, Pond GR, Crowson CS, O'Fallon WM, Gabriel SE. Trends in incidence and mortality in rheumatoid arthritis in Rochester, Minnesota, over a forty-year period. Arthritis Rheum 2002 Mar;46(3):625-31.

(168) Myasoedova E, Crowson CS, Kremers HM, Therneau TM, Gabriel SE. Is the incidence of rheumatoid arthritis rising?: results from Olmsted County, Minnesota, 1955-2007. Arthritis Rheum 2010 Jun;62(6):1576-82.

(169) Cross M, Smith E, Hoy D, Carmona L, Wolfe F, Vos T, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis 2014 Jul;73(7):1316-22.

(170) Hecht C, Englbrecht M, Rech J, Schmidt S, Araujo E, Engelke K, et al. Additive effect of anti-citrullinated protein antibodies and rheumatoid factor on bone erosions in patients with RA. Ann Rheum Dis 2014 Aug 12.

(171) Sokolove J, Johnson DS, Lahey LJ, Wagner CA, Cheng D, Thiele GM, et al. Rheumatoid factor as a potentiator of anti-citrullinated protein antibody-mediated inflammation in rheumatoid arthritis. Arthritis Rheumatol 2014 Apr;66(4):813-21.

(172) Radovits BJ, Fransen J, Al SS, Eijsbouts AM, van Riel PL, Laan RF. Excess mortality emerges after 10 years in an inception cohort of early rheumatoid arthritis. Arthritis Care Res (Hoboken) 2010 Mar;62(3):362-70.

(173) Widdifield J, Bernatsky S, Paterson JM, Tomlinson G, Tu K, Kuriya B, et al. Trends in excess mortality among patients with rheumatoid arthritis in ontario, Canada. Arthritis Care Res (Hoboken) 2015 Jan 26.

(174) Dadoun S, Zeboulon-Ktorza N, Combescure C, Elhai M, Rozenberg S, Gossec L, et al. Mortality in rheumatoid arthritis over the last fifty years: Systematic review and meta-analysis. Joint Bone Spine 2012 Apr;(0).

(175) Gonzalez A, Maradit KH, Crowson CS, Nicola PJ, Davis JM, III, Therneau TM, et al. The widening mortality gap between rheumatoid arthritis patients and the general population. Arthritis Rheum 2007 Nov;56(11):3583-7.

(176) Norton S, Carpenter L, Nikiphorou E Walsh D, Kiely P, Dixey J, Young A. Excess mortality in rheumatoid arthritis: gains in life expectancy over 25 years. Rheumatology 53[Supplement 1], i43-i44. 2014.
Ref Type: Abstract

(177) Solomon DH, Karlson EW, Rimm EB, Cannuscio CC, Mandl LA, Manson JE, et al. Cardiovascular morbidity and mortality in women diagnosed with rheumatoid arthritis. Circulation 2003 Mar 11;107(9):1303-7.

(178) Finckh A, Choi HK, Wolfe F. Progression of radiographic joint damage in different eras: trends towards milder disease in rheumatoid arthritis are attributable to improved treatment. Ann Rheum Dis 2006 Sep;65(9):1192-7.

(179) Welsing PM, Fransen J, van Riel PL. Is the disease course of rheumatoid arthritis becoming milder? Time trends since 1985 in an inception cohort of early rheumatoid arthritis. Arthritis Rheum 2005 Sep;52(9):2616-24.

(180) Gan RW, Trouw LA, Shi J, Toes RE, Huizinga TW, Demoruelle MK, et al. Anti-carbamylated Protein Antibodies Are Present Prior to Rheumatoid Arthritis and Are Associated with Its Future Diagnosis. J Rheumatol 2015 Jan 15.

(181) Alarcon GS, Koopman WJ, Acton RT, Barger BO. Seronegative rheumatoid arthritis. A distinct immunogenetic disease? Arthritis Rheum 1982 May;25(5):502-7.

(182) Panayi GS, Celinska E, Emery P, Griffin J, Welsh KI, Grahame R, et al. Seronegative and seropositive rheumatoid arthritis: similar diseases. Br J Rheumatol 1987 Jun;26(3):172-80.

(183) Daha NA, Toes RE. Rheumatoid arthritis: Are ACPA-positive and ACPA-negative RA the same disease? Nat Rev Rheumatol 2011 Apr;7(4):202-3.

(184) Padyukov L, Seielstad M, Ong RT, Ding B, Ronnelid J, Seddighzadeh M, et al. A genome-wide association study suggests contrasting associations in ACPA-positive versus ACPA-negative rheumatoid arthritis. Ann Rheum Dis 2011 Feb;70(2):259-65.

(185) RA-MAP Consortium. The TACERA study; a longitudinal observational study of patients with early rheumatoid arthritis. http://www gsttbrc com/tacera/about aspx?lang=1 2012 May 29 [cited 2015 Mar 11];

(186) Morgan C, Lunt M, Bunn D, Scott DG, Symmons DP. Five-year outcome of a primary-care-based inception cohort of patients with inflammatory polyarthritis plus psoriasis. Rheumatology (Oxford) 2007 Dec;46(12):1819-23.

(187) Stavropoulos-Kalinoglou A, Metsios GS, Koutedakis Y, Kitas GD. Obesity in rheumatoid arthritis. Rheumatology (Oxford) 2011 Mar;50(3):450-62.

(188) Boyle DL, Soma K, Hodge J, Kavanaugh A, Mandel D, Mease P, et al. The JAK inhibitor tofacitinib suppresses synovial JAK1-STAT signalling in rheumatoid arthritis. Ann Rheum Dis 2014 Nov 14.

Appendix 1: Embase and Medline search strategies

OVID Medline

	Searches	Results
1	rheumatoid arthritis.mp. or exp rheumatoid arthritis/	113359
2	limit 1 to yr="2010 -Current	18803
3	classification/ or disease classification/ or classification criteria.mp.	11222
4	1 and 2 and 3	263

OVID Embase

	Searches	Results
1	rheumatoid arthritis.mp. or rheumatoid arthritis/	161895
2	limit 1 to (full text and english language and yr="2010 –Current)	7877
3	disease classification/ or classification/ or classification criteria.mp.	350870
4	1 and 2 and 3	293

Appendix 2: NOAR Baseline Questionnaires

NOAR Number: _____ DOB:

A. Demographic Details

Ethnic origin _____

Place of birth _____

Did the patient develop their joint problems while living in Norfolk?

Yes ☐ No ☐

Has the patient completed an EPIC form?

Yes ☐ No ☐

B. Occupational History

Occupation: ☐

1 Working now ☐
2 Unemployed ☐
3 Retired ☐
4 Off sick ☐
5 Never worked ☐
6 Housewife/mother/carer ☐

If 2, 3 or 4 date last worked d d m m y y ☐☐☐☐☐☐

Is (was) the patient an employee ☐
Self-employed ☐
Manager ☐
Foreman/supervisor ☐
Other employee ☐ How many people are supervised ☐☐

Employer's name:
_____

Patient's husband

Occupation: ☐

Is (was) he an employee ☐
Self-employed ☐
Manager ☐
Foreman/supervisor ☐
Other employee ☐ How many people are supervised ☐☐

Yes No

Has the patient been off work in the last 12 months? ☐ ☐

Total number of days off ☐☐☐
How many due to joints? ☐☐☐

C. Comorbidity

1. Has the patient ever had any of the following conditions:
(GP or nurse diagnosis)

	Yes	No	Yr of onset	
Angina				
High blood pressure				
Heart attack				
Heart failure				
Stroke				
TIA (mini-stroke/transient visual loss)				
Diabetes				
Stomach Ulcer				
Liver disease				
Kidney failure				
Cancer (except skin cancer)				
Psoriasis				
Depression				
Emphysema, chronic bronchitis or asthma				
Glaucoma				

2. If the patient is currently diabetic, is (s)he
Insulin ☐ Tablet ☐ Diet ☐ dependent ☐ Controlled ☐ controlled ☐

3. If the patient has had cancer, please specify the primary site

_____

Other Intercurrent Illness

4. Include other chronic conditions (not listed in C1) and date of onset.

_____

_____

_____

_____

_____

_____

D. Family History

Does the patient have a family history (first degree relative) with

	Yes	No
RA		
Psoriasis		
MI (if yes, include age & sex of relative)		
Premature cardiovascular death (if yes, include age & sex of relative) (males aged < 55, females aged <65 years)		

E. Smoking History

Has the patient ever smoked as much as one cigarette a day for as long as a year? (Yes / No)

(If NO, got to Section F)

If YES, how old was (s)he when (s)he started smoking regularly (Years)

Did (s)he smoke at the following ages? If so, how many cigarettes a day?

			Yes	No			N/A
Age	20	Smoker			Cigarettes per day		
Age	30	Smoker			Cigarettes per day		
Age	40	Smoker			Cigarettes per day		
Age	50	Smoker			Cigarettes per day		
Age	60	Smoker			Cigarettes per day		
Age	70	Smoker			Cigarettes per day		

Does (s)he smoke now? (Yes / No)

If YES, how many cigarettes each day

If (s)he has stopped smoking, how old was (s)he when (s)he gave up

F. Reproductive history

		Yes	No
1.	Has the patient ever been pregnant?		
2.	How many pregnancies has she had?		
	Years of - live births		
	- still births		
	- miscarriages		
	- terminations		
3.	Oral contraceptive pill:	Yes	No
	- now		
	- ever (for 3/12)		
4.	Year of menopause		
	Or		
5.	Year of hysterectomy		
6.	Age at menarche, if under 16		
7.	Has the patient had HRT?		
	- now		
	- ever		

If YES, for how may months (in total) has the patient taken HRT?

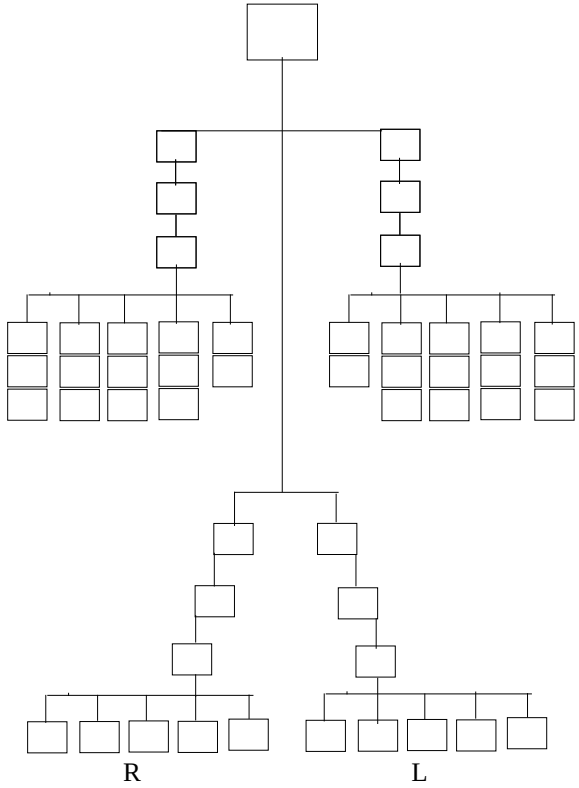
G. History of joint symptoms

1 Duration of morning stiffness

Maximal Today

						Minutes

2. Which joints were painful (P) and/or swollen (S) in the first two weeks (1) or ever (2)



H. Medication

1. Current medication (list all drugs)

_____ _____ _____

_____ _____ _____

2. DMARD Therapy

Since disease onset has the patient received:

	Date Started	Date Stopped	Reason For Stopping	Dose at Time Stopped
IM Gold				
Auranofin				
Penicillamine				
Sulphasalazine				
Chlor/HCQ				
Methotrexate				
Azathioprine				
Cyclophosphamide				
Cyclosporin				
Steroids				
Leflunomide				
Other				

Code for stopping:
1. Adverse reaction a) skin, b) blood, c) gut, d) renal, e) other
2. Inefficacy
3. Disease remission
4. Planned course complete
5. Lack of compliance
6. Other

If on steroid since symptom onset, state maximum dose:

3. BIOLOGIC THERAPY: since disease onset, has the patient received:

	Date Started	Date Stopped	Reason For Stopping	Dose at Time Stopped
Infliximab				
Etanercept				
Anakinra				
Adalimumab				
Other (specify				

Code for stopping:
1. Adverse reaction a) skin, b) blood, c) gut, d) renal, e) other
2 Inefficacy
3. Disease remission
4. Planned course complete
5. Lack of compliance
6. Other

4. STEROID THERAPY since symptom onset

	Yes	No
IM steroid?		
How many doses?		
How many joint injections since symptom onset?		

5. Statins

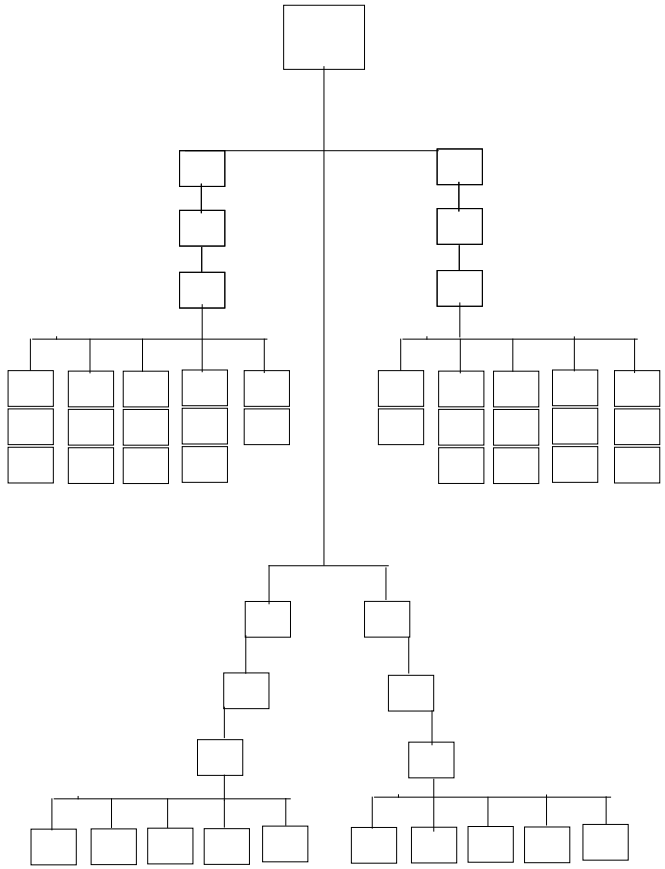
6. NSAID Therapy since symptom onset

Please record all the NSAIDs which the patient has taken **since symptom onset** (together with start and stop dates where possible)

Drug (NSAID)	Date Started	Date Stopped

I. **<u>Examination</u>**

a) Joint activity
Which joints are tender (T), swollen (S) or both (B) now?



R L

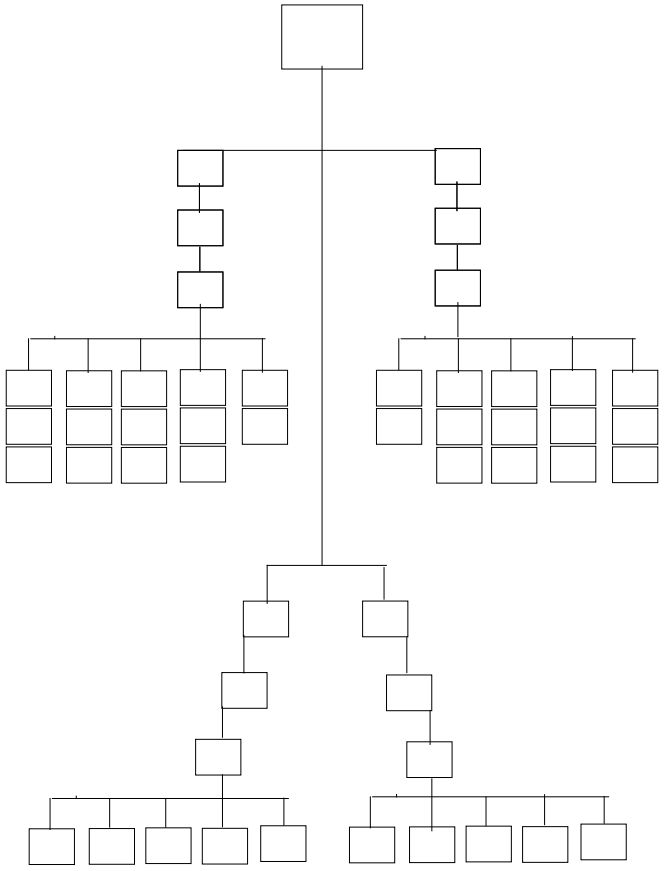
Does the patient have:

	Yes	No
Nodules		
Dry eyes		
Leg ulcers		
Psoriasis		

Left / Right Handed?

b) Joint deformity

Which joints are deformed (D) or have been operated on (O) now



R L

c) HEIGHT (cm) …….

d) WEIGHT (kg) ……..

e) BP — Systolic (R) ☐☐☐ Diastolic (R) ☐☐☐ Systolic (L) ☐☐☐ Diastolic (L) ☐☐☐

Notify GP if BP > 140/90 ☐☐

Pulse ☐☐☐

Ask GP to review soon if > 160/100 ☐☐

f) Waist Circumference ☐☐☐ cm

g) Hip Circumference ☐☐☐ cm

h) Most Recent ESR ☐☐☐ Date ☐☐☐☐☐☐

i) Most recent CRP ☐☐☐ Date ☐☐☐☐☐☐

J. <u>Hospital attendance for arthritis: since symptom onset</u>

1. Has the patient been referred to hospital for arthritis since symptom onset. Yes ☐ No ☐

2. Date of 1st appointment

d	d	m	m	y	Y

3. How many times has the patient been admitted to hospital for arthritis since symptom onset? ☐☐

4. Any joint surgery since symptom onset? Yes ☐ No ☐

5. Joint:_____
Operation: _____
Date:_____

Name of Consultant?

_____

Rheumatologist / Orthopaedic Surgeon / Other (state)

K. Summary

1.	Morning stiffness more than one hour	
2.	Arthritis of 3 or more joint areas (PIP, MCP, wrist, elbow, knee, ankle, MTP)	
3.	Arthritis of hand joint (wrist, MCPs, PIP)	
4.	Symmetry	
5.	Rheumatoid nodules	
6.	ARA criteria satisfied	
7.	No active arthritis on assessment	
8.	Register: Inflammatory arthritis (I) Rheumatoid arthritis (R) NOAR Criteria not met (N)	

L. Administration

	Yes	No
1. Has the CLINHAQ been collected?		

2. Is the patient suitable for cardiovascular study:

	Yes	No
Age 18-64		
Symptom onset < 24 months		
Consent		

If YES, make arrangements for blood to be taken fasting.
Initiate arrangements for Doppler scan.

3. Blood taken

	Yes	No
Serum		
EDTA		

Box number (Clear top) | | | |

(Green top) | | | |

	Yes	No
Blood suitable for homocysteine testing (ie, collected and frozen within 1 hour?		

If No blood, arrangements made:

Signature of Nurse

	d	d	m	m	y	y
Date						

Dry Eyes:	Yes	No
Has the patient had daily, persistent, troublesome dry eyes for more than three months?		
Do they have a sensation of sand or gravel in their eye?		
Do they need to use eye drops containing tear substitutes more than three times a day?		
Dry Mouth		
Has the patient had a daily feeling of dry mouth for more than three months?		
Do they keep getting swollen salivary glands (located between the jaw and ears)?		
Do they frequently drink liquids to help them swallow food?		

Clinical Health Assessment Questionnaire (CLINHAQ)

(All Anniversaries)

NOAR ID _____ DOB _____ Date _____

We are interested in learning how your illness affects your ability to function in daily life.

Please tick the response which best describes your usual abilities OVER THE PAST WEEK:

	Without any difficulty	With some difficulty	With much difficulty	Unable to do	For Office use
DRESSING & GROOMING: Are you able to:					
Dress yourself, including tying shoelaces & doing buttons?	_____	_____	_____	_____	
Shampoo your hair?	_____	_____	_____	_____	Dressing ____
RISING: Are you able to:					
Stand up from an armless straight chair?	_____	_____	_____	_____	
Get in and out of bed?	_____	_____	_____	_____	Rising ____
EATING: Are you able to:					
Cut your meat?	_____	_____	_____	_____	
Lift a full cup or glass to your mouth?	_____	_____	_____	_____	Eating
Open a new carton of milk (or soap powder)?	_____	_____	_____	_____	____
WALKING: Are you able to:					
Walk outdoors on flat ground?	_____	_____	_____	_____	Walking
Climb up five steps?	_____	_____	_____	_____	____

Please tick any AIDS or DEVICES that you usually use for any of these activities:

_____ Cane (W) _____ Walking frame (W) _____ Built up or special utensils

_____ Crutches (W) _____ Wheelchair (W) _____ Special or built up chair

_____ Devices used for dressing (button hook, zipper pull, long handled shoe horn)

_____ Other, specify _____

Please tick any category for which you usually need HELP FROM ANOTHER PERSON:

_____ Dressing & grooming _____ Eating

_____ Rising _____ Walking

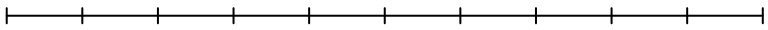
We are also interested in learning whether or not you are affected by pain because of your illness.

How much pain have you had because of your arthritis IN THE PAST WEEK?

Place a mark on the line to indicate the severity of the pain

Painscale ___

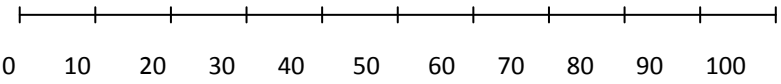
No pain Severe pain



0 10 20 30 40 50 60 70 80 90 100

How much trouble have you had with your stomach (ie nausea, heartburn, bloating, pain, etc) IN THE PAST WEEK? Place a mark on the line below to indicate the severity of the problem:

No stomach problem Severe stomach problem



Please tick the response which best describes your usual abilities OVER THE PAST WEEK:

	Without any difficulty	With some difficulty	With much difficulty	Unable to do
HYGIENE: Are you able to:				
Wash and dry your entire body?	_____	_____	_____	_____
Take a bath?	_____	_____	_____	_____
Get on and off the toilet?	_____	_____	_____	_____
REACH: Are you able to:				
Reach and get down a 5lb object (eg a bag of potatoes) from just above your head?	_____	_____	_____	_____
Bend down to pick up clothing from the floor?	_____	_____	_____	_____
GRIP: Are you able to:				
Open car doors?	_____	_____	_____	_____
Open jars which have been previously opened?	_____	_____	_____	_____
Turn taps on and off?	_____	_____	_____	_____
ACTIVITIES: Are you able to:				
Run errands and shop?	_____	_____	_____	_____
Get in and out of a car?	_____	_____	_____	_____
Do chores (vacuuming, housework or light gardening)?	_____	_____	_____	_____

For Office Use

Hygiene ____

Reach _____

Grip _____

Activity ____

Please tick any AIDS OR DEVICES that you usually use for any of these activities:

_____ Bath rail (H) _____ Bath seat (H)

_____ Raised toilet seat (H) _____ Jar opener (for jars previously opened) (G)

_____ Long-handled appliances for reach (R) ____Other (specify) _____

Please tick any categories for which you usually need HELP FROM ANOTHER PERSON:

_____ Hygiene _____ Gripping and opening things

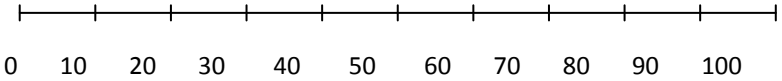
_____ Reach _____ Errands and chores

In general, would you say that your HEALTH IS:

_____ Excellent _____ Good _____ Fair _____Poor

Consider ALL THE WAYS THAT YOUR ARTHRITIS AFFECTS YOU, RATE HOW YOU ARE DOING on the following scale by placing a mark on the line below:

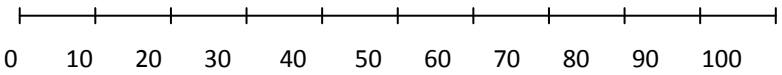
Very well Very poorly



We are interested in knowing about any problems that you may have been having with fatigue.

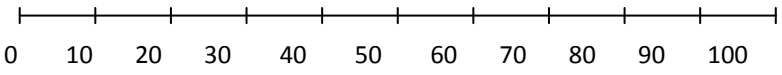
How much of a problem has fatigue or tiredness been for you IN THE PAST WEEK? Place a mark on the line below:

Fatigue is no problem Fatigue is a major problem



How much of a problem has sleep (ie resting at night) been for you IN THE PAST WEEK? Place a mark on the line below:

Sleep is no problem Sleep is a major problem



Please tick the most appropriate answer for each Question. Try to answer every question.

	ALWAYS	VERY OFTEN	SOMETIMES	ALMOST NEVER	NEVER
1. During the PAST MONTH, how often have you felt tense or "highly strung"?	_____	_____	_____	_____	_____
2. How often have you been bothered by nervousness or your "nerves" during the past month?	_____	_____	_____	_____	_____
3. How often during the PAST MONTH were you able to relax without difficulty?	_____	_____	_____	_____	_____
4. How often during the PAST MONTH have you felt relaxed and free of tension?	_____	_____	_____	_____	_____
5. How often during the PAST MONTH have you felt calm and peaceful?	_____	_____	_____	_____	_____
6. During the PAST MONTH, how often have you enjoyed the things you do?	_____	_____	_____	_____	_____
7. During the PAST MONTH, how often have you been in low or very low spirits?	_____	_____	_____	_____	_____
8. How often during the PAST MONTH did you feel that nothing turned out the way you wanted it to?	_____	_____	_____	_____	_____
9. During the PAST MONTH, how often did you feel that others would be better off if you were dead?	_____	_____	_____	_____	_____
10. How often during the PAST MONTH, did you feel down in the dumps that nothing could cheer you up?	_____	_____	_____	_____	_____

The statements below concern your personal beliefs. PLEASE CIRCLE THE NUMBER beside each statement that best describes how you feel about the statement.

	STRONGLY DISAGREE	DISAGREE	DO NOT AGREE OR DISAGREE	AGREE	STRONGLY AGREE
My condition is controlling my life	1	2	3	4	5
I would feel helpless if I couldn't rely on other people for help with my condition	1	2	3	4	5
No matter what I do, or how hard I try, I just can't seem to get relief from my pain	1	2	3	4	5
I am not coping effectively with my condition	1	2	3	4	5
It seems as though fate and other factors beyond my control affect my condition	1	2	3	4	5

How satisfied are you with your HEALTH NOW?

_____Very satisfied
_____ Somewhat satisfied
_____ Neither satisfied or dissatisfied
_____ Somewhat dissatisfied
_____ Very dissatisfied

London School of Hygiene Chest Pain Questionnaire

PART A.

a. Have you ever had any pain or discomfort in your chest?

1. ☐ Yes 2. ☐ **If No**

(Go to Part C)

b. Do you get this pain or discomfort when you walk uphill or hurry?

1. ☐ Yes 2. ☐ **If No**

(Go to Part B)

c. Do you get it when you walk at an ordinary pace on the level?

1. ☐ Yes 2. ☐ No

d. When you get any pain or discomfort in your chest what do you do?

1. ☐ Stop
2. ☐ Slow down
3. ☐ Continue at the same pace

e. Does it go away when you stand still?

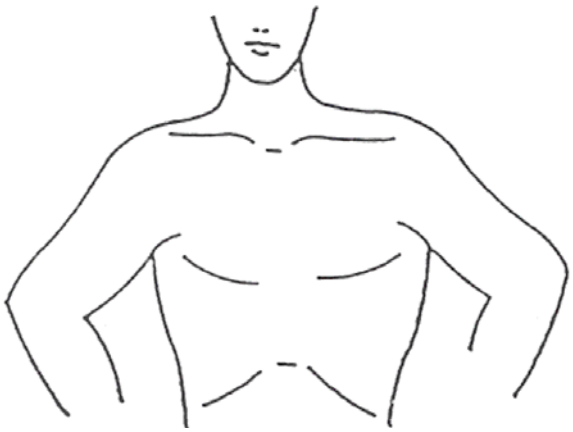
1. ☐ Yes 2. ☐ No

f. How soon?

1. ☐ 10 minutes or less
2. ☐ More than 10 minutes

g. When do you get this pain or discomfort?

Mark the place(s) with X on the diagram



PART B.

Have you ever had a severe pain across the front of your chest lasting for half an hour or more?

1. ☐ Yes 2. ☐ No

PART C.

a. Do you get a pain in either leg on walking?

1. ☐ Yes (Complete remaining questions)
2. ☐ No (No need to answer questions below)

b. Does this pain ever begin when you are standing still or sitting?

1. ☐ Yes 2. ☐ No

c. Do you get this pain in your calf(or calves)?

1. ☐ Yes 2. ☐ No

d. Do you get it when you walk uphill or hurry?

1. ☐ Yes 2. ☐ No

e. Do you get it when you walk at an ordinary pace on the level?

1. ☐ Yes 2. ☐ No

f. Does the pain ever disappear while you are still walking?

1. ☐ Yes 2. ☐ No

g. What do you do if you get it when you are walking?

1. ☐ Stop
2. ☐ Slow down
3. ☐ Continue at the same pace

h. What happens to it if you stand still?

1. ☐ Usually continues more than 10 minutes.
2. ☐ Usually disappears in 10 minutes or less.

Appendix 3: NOAR follow up questionnaire

NOAR Number:_____DOB:_____

1st, 2nd, 3rd ,5th & 7th Anniversary

Occupational History

		Yes	No
1.	Has there been any change in the patient's employment since last seen? If NO, proceed to question 5.		

2. If YES, is the patient currently

a	Working		f	Never worked	
b	Unemployed		g	Housewife/mother/carer	
c	Off sick		h	Student	
d	Retired		i	Working reduced hours	
e	Not working – health grounds		i	Working, but there has been a change in duties	

3. If b, c, d, e, f, g, date last worked

d	d	m	m	y	y

4. If a, new occupation: _____

5 If e, i or j give details _____

Is the patient

an employee	
self-employed	
manager	
foreman/supervisor	
other employee	

How many people are supervised _____

	Yes	No
5. Has the patient been off work since last seen?		

Total number of days off			
How many due to joints?			

B. Co-Morbidity

4. Has the patient developed any of the following conditions since last seen?

	Yes	No	Hospital Admission or Referral
Angina			
High blood pressure			
Heart attack			
Heart failure			
Stroke			
TIA			
Diabetes			
Stomach Ulcer			
Liver disease			
Kidney failure			
Cancer (except skin cancer)			
Psoriasis			
Depression			

Emphysema, chronic bronchitis or asthma			
Glaucoma			

5. If the patient is currently diabetic, is (s)he

Insulin dependent		Tablet controlled		Diet controlled	

6. If the patient has had cancer, please specify the primary site

_____

C. Smoking

	Yes	No
Does the patient currently smoke cigarettes?		
How many cigarettes does (s)he smoke per day		

Co-Morbidity (cont)

4. Other Intercurrent Illness

Include other chronic conditions, hospital admissions and surgery (not listed in B1) and date of onset.

_____

_____

_____

_____

_____

_____

_____

D. Reproductive history

	Yes	No
Has the patient been pregnant since last seen?		

Outcome	d	d	m	m	y	y
Termination						
Miscarriage						
Stillbirth						
Livebirth						

	Yes	No
Has the patient passed through the menopause (no periods for six months) since last seen?		
Had a hysterectomy since last seen?		
If yes, Abdominal or vaginal?		
Since last seen, has the patient taken:		
the pill		
HRT		

If on HRT, total number of months

E. Joint symptoms

	Swelling		EMS > 30 minutes	
	Yes	No	Yes	No
a) Have you had any joint swelling or stiffness in the last 6 months?				

b) How long were you stiff for this morning? Minutes

F. Medication

2. Current medication (list all drugs)

_____

_____

_____

_____

2. DMARD Therapy

Since the last assessment, has the patient received:

	Date Started	Date Stopped	Reason For Stopping	Dose at Time Stopped
IM Gold				
Auranofin				
Penicillamine				
Sulfasalazine				
Chlor/HCQ				
Methotrexate				
Azathioprine				
Cyclophosphamide				
Ciclosporin				
Steroids				

Leflunomide																							
Other																							

Code for stopping:

1. Adverse reaction a) skin, b) blood, c) gut, d) renal, e) other
2. Inefficacy
3. Disease remission
4. Planned course complete
5. Lack of compliance
6. Other

If on steroid since last assessment, state maximum dose | | | . | | | mg

3. BIOLOGIC THERAPY:

since last assessment, has the patient received:

	Date Started								Date Stopped								Reason For Stopping			Dose at Time Stopped
Infliximab																				
Etanercept																				
Anakinra																				
Adalimumab																				
Other (specify																				

Code for stopping:

1. Adverse reaction a) skin, b) blood, c) gut, d) renal, e) other
2. Inefficacy
3. Disease remission
4. Planned course complete
5. Lack of compliance
6. Other

4. STEROID THERAPY since last assessment

	Yes	No
IM steroid?		
How many doses?		
How many joint injections since last seen?		

5. STATINS

	Now	No
Is the patient on a statin?		

If YES, for how may months (in total) has the patient taken a statin?

G. Examination

Height		cm
Weight		kg
Waist circumference		cm
Hip circumference		cm
Systolic BP (Right)		
Diastolic BP (Right)		

Pulse rate

Systolic BP (Left)	mm Hg
Diastolic BP (Left)	mm Hg

From patient records, Latest ESR & CRP

ESR

	d d m m y y
mm	

CRP

	d d m m y y

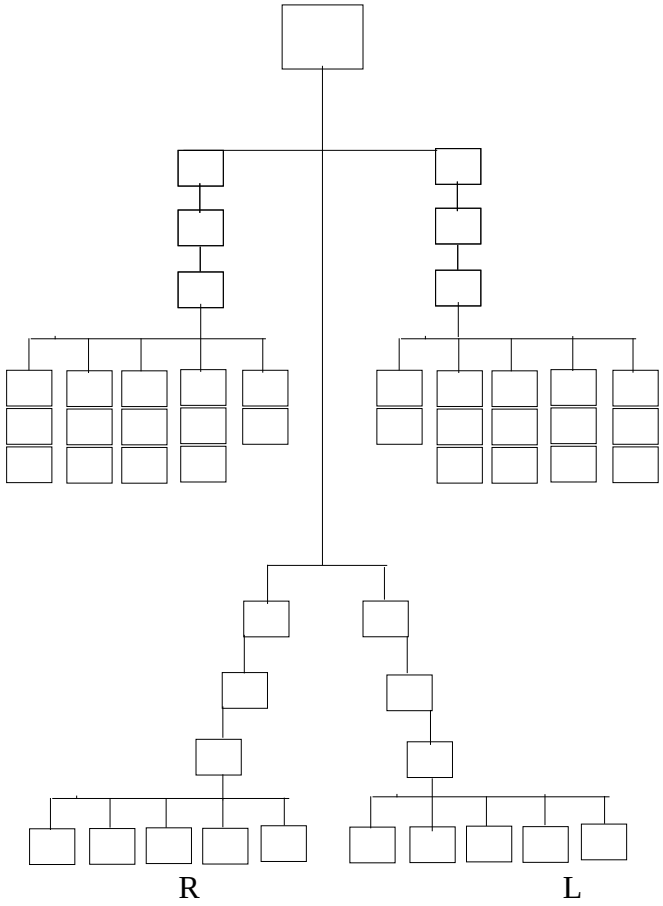
For Norwich Use Only

BP > 140/90
GP Informed? Yes / No

BP > 160/100
GP informed for review soon?

H. Joint Examination

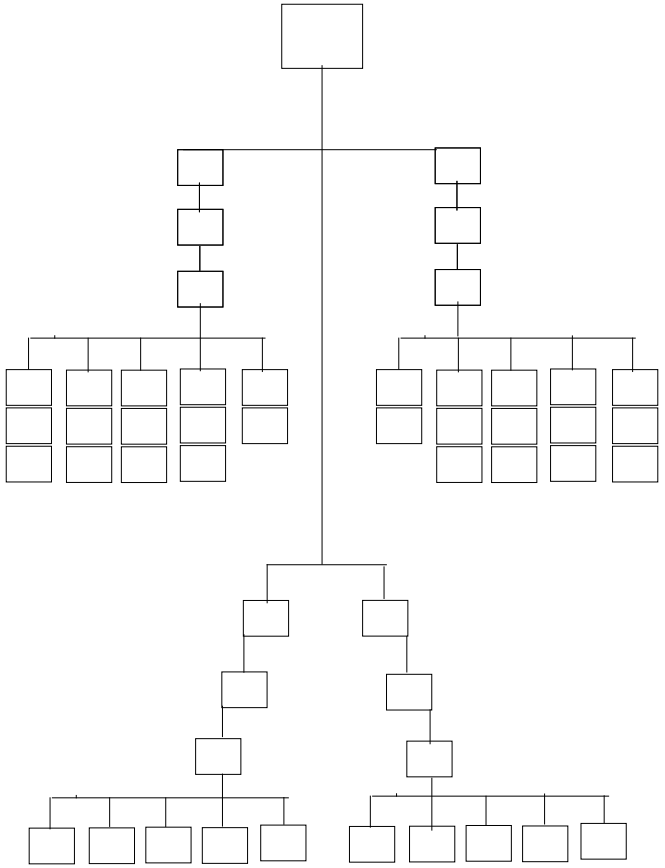
a) Joint activity
Which joints are tender (T), swollen (S) or both (B) now



Does the patient have:	Yes	No
Nodules		
Dry eyes		
Leg ulcers		
Psoriasis		

b) Joint deformity

Which joints are deformed (D) or have been operated on (O) now



R L

I. Hospital Attendance for arthritis

	Yes	No
1. Has the patient been referred to hospital for arthritis in the last year?		

	d	d	m	m	y	y
2. If yes, date of 1st appointment?						

3. How often has patient attended hospital in last year?		
a). Doctor		
b). Nurse		
c). Day case		

4. How many times has patient been admitted to hospital for arthritis in the last year?		

	Yes	No
5. Any joint surgery in the last 12 months?		

6. Details:

Joint:_____
_____

Operation:
_____

Date:
_____

ARA criteria at end of year ~~one~~
(Arthritis = S or B at 1st & 2nd anniv, but 'B' only at 3rd

1. Morning stiffness more than one hour		
2. Arthritis of 3 or more joint areas (PIP, MCP, wrist, elbow, knee, ankle, MTPP)		
3. Arthritis of hand joints (wrists, MCP, PIP)		
4. Symmetry		
5. Rheumatoid nodules		

Summary (resolution) criteria.

If you Agree with the statement, please tick the box beside it.

(Arthritis = S or B at 1st and 2nd anniv, but 'B' only at 3rd)

1. No arthritis on examination	
2. No symptoms of swelling or stiffness (> 30 mins) for more than 6 months	
3. Not on second-line drug or steroids in the last 3 months	

Register at end of this year:

Rheumatoid (R)	
Rheumatoid Resolved (RAR)	
Inflammatory (I)	
Inflammatory Resolved (IAR)	
Other (state which)	

X-ray: Arrange an X-ray at year one (1st anniversary) if patient in cardiovascular study

Yes		Not required	
No		Declined	

Blood
From 2005: bloods required from *everyone*.

CLINHAQ collected? Yes / No

Signature of Research Nurse: _____ Date:
_____

	Yes	No
Dry Eyes: Has the patient had daily, persistent, troublesome dry eyes for more than three months?		
Do they have a sensation of sand or gravel in their eye?		
Do they need to use eye drops containing tear substitutes more than three times a day?		
Dry Mouth		
Has the patient had a daily feeling of dry mouth for more than three months?		
Do they keep getting swollen salivary glands (located between the jaw and ears)?		
Do they frequently drink liquids to help them swallow food?		

Appendix 4: Letter to the custodians of longitudinal observational studies of inflammatory arthritis

Dear

My name is Jenny Humphreys, I am a Clinical Research Fellow at the University of Manchester. My PhD project is investigating epidemiological issues relating to the 2010 ACR/EULAR classification criteria for RA.

Currently, I and my supervisory team (Professor Deborah Symmons, Dr Kimme Hyrich and Dr Suzan Verstappen) are debating how to apply the criteria in patients with established disease. Our discussions have highlighted a number of potential challenges. For example, patients whose disease is in remission because of DMARD therapy may not satisfy the criteria and, if their disease has been well controlled throughout, may not have erosions. In addition we questioned the usefulness of an isolated acute phase reactant measure in established disease.

In part this has been addressed by the work presented at EULAR, defining what is meant by 'erosive disease' in the new criteria, by Professor van der Heijde (Ann Rheum Dis 2012;71(Suppl3):25). To explore this further, we analysed a dataset from a cross-sectional prevalence survey conducted in 2000 in Norfolk, UK (Rheumatology 2002;41(7):793-800), applying both the erosion rules as well as the 2010 classification criteria to estimate prevalence. There were missing data on some of the variables. Nevertheless, in a complete case analysis we found a proportion of the patients could be classified as RA by the 1987 criteria, but were not captured using either of the above methods. We feel this is an important issue because the 1987 criteria have usually been accepted as fit for purpose in identifying patients with established disease.

We are interested in viewpoints of others involved in observational cohorts on this subject, and have developed a short questionnaire (overleaf), we would be very grateful if you would complete and return via email to:
jennifer.humphreys@manchester.ac.uk.

Please note that any replies given or views expressed would be treated as entirely confidential and would not be used in any publications without first consulting with yourselves.

Questionnaire

1. Are there any specific challenges you think may arise in applying the 2010 criteria in established disease (e.g. patients on DMARDs, missing data)?

2. Do you think the criteria could be adapted to be more applicable in established disease and if so, how?

3. Have you had any experience applying the 2010 criteria in established disease? If so, what were the challenges and how did you deal with them?

4. Are the variables required to apply the 2010 criteria available in observational cohorts you are involved with?

5. Any other comments
6.