UNIVERSITY OF GOTHENBURG

SAHLGRENSKA ACADEMY

How can we identify misdiagnosis in rheumatoid arthritis?

Degree Project in Medicine

Sophie Hallberg

Programme in Medicine

Gothenburg, Sweden 2021

Supervisor: PhD Monica Leu Agelii

Dept of Rheumatology and Inflammation Research,

Göteborg University

Index

Abbreviations

ACR - American College of Rheumatology

Anti-CCP - Anti-Cyclic Citrullinated Peptide

BARFOT - Better Anti Rheumatic Pharmaco Therapy

CRP - C-reactive protein

DAS28 - Disease Activity Score

DMARDs - Disease Modifying Antirheumatic Drugs

ES - erosion score

ESR - Erythrocyte sedimentation rate

EULAR - European League against Rheumatism

GC - Glucocorticosteroids

JNS - joint narrowing score

MTX - Methotrexate

RA - Rheumatoid Arthritis

RF - Rheumatoid Factor

SAL – Sulphasalazine

SHS - Sharp van der Heijde score

VAS - Visual Analog Scale

UN - Undifferentiated arthritis

Abstract

Background: Rheumatoid arthritis (RA) is a common systemic autoimmune disease, with largely unknown pathogenesis, mainly affecting joints, leading to deformities with loss of function. The prevalence is 0.5 – 1.0 % in the general population. Diagnosis in an early stage of the disease is difficult but essential for suppressing inflammation and preventing damage and deformities in joints. As diagnosis criteria for RA are missing, the diagnosis is often established based on classification criteria.

Objectives: This study examines differences between actual and misdiagnosed in a long-term cohort where patients were included based on the ACR1987 classification criteria for rheumatoid arthritis.

Methods: Of the 2541 patients from the BARFOT (Better AntiRheumatic Pharmaco Therapy) cohort, the RA diagnosis was changed in 44 patients (RA-change group). This group is compared patients who kept their RA (RA-keep group). The BARFOT cohort was followed for 15 years. At inclusion all patients fulfilled the ACR 1987 classification criteria for RA. The two groups were compared during their first two years in cohort regarding the number and type of classification criteria as well disease activity, medication, and radiographic changes.

Results: Half of the RA-change group were classified as RA 5 years after diagnosis. The RA-keep group had a higher proportion of RF-positivity (63.1% vs 21.4% in RA-change group, p=0.001) and was more likely to fulfill >4 ACR1987 criteria (63.5% vs 34.1%, p=0.001). There was a higher proportion of patients with radiographic joint destruction at inclusion in RA-keep (26.5%) vs RA-change (12.2%, p=0.04). The erythrocyte sedimentation rate was increased in the RA-keep compared to the RA-change group over 2 years from diagnosis (p=0.02).

Conclusions: Diagnosis of RA should be reconsidered for patients who are RF-negative and do not fulfil more than 4 ACR1987 criteria.

Keywords: Rheumatoid arthritis, Misdiagnosis, Inflammatory disease, ACR1987-criteria, Rheumatoid factor, DAS28

Background

Rheumatoid arthritis (RA) is a systemic chronic inflammatory disease that mainly affects the joints, but also other tissues [1] and can lead to lifelong chronic disability. The prevalence is 0.5 – 1.0 % in the general population and the disease is up to three times more common in women compared to men[1-3]. The pathogenesis of RA is largely unknown, but the current evidence points to a multifactorial aetiology, with age, female gender, genetic factors and smoking among the risk factors [4]. Additionally, exposure to environmental factors such as silica dust seems to contribute to the risk of developing RA [5].

Symptoms

The most common symptoms are due to inflammation and includes pain, swelling and stiffness of the small joints in the hands and feet [4]. These symptoms are particularly apparent in the morning. Around 90 % of RA patients are affected by symptoms from the hands, which, if not treated, lead to deformities, joint damage, and loss of function such as to grip, pinch, grasp and overall limit the patient's movement [6]. As well as symptoms from other joint-areas, RA involve a wide range of signs such as fatigue, fever, pulmonary involvement, vasculitis, rheumatoid nodules and loss of physical function. RA affects the whole body and also puts patients at a higher risk for cardiovascular disease, serious infections and increased mortality [1, 4]. In contrast to the general population, people with RA have a higher risk for anxiety, sleep deprivation, depression, chronic pain and reduced capacity both at work and in social situations [7, 8]. This results in a diminished quality of life for the patients and puts a substantial burden on society as a whole both because of a reduced work capacity and high costs for medical treatment and social support [9].

Treatment

The available treatment aims at suppressing inflammation. Even if it is not curative, the prognosis of most individuals with RA has dramatically improved over the last couple of decades: from having to manage several of the aforementioned comorbidities to halting disease progression and even going into remission. In one meta-analysis including thirty-one studies with 82 450 RA patients in total, where 17% of the patients after 3 months, and 23% after 24 months showed remission [10]. Some reasons for this improvement are earlier disease recognition and treatment, effective use of disease modifying anti rheumatics drugs (DMARDs), and a treat-to-target strategy where the target for treatment is remission or a low disease activity score, which can be achieved with tight monitoring and change of treatment where the target is not met [4]. Disease activity score 28 (DAS28) is a summary score representing disease status and is calculated based on the number of swollen and tender joints out of 28 included joints, erythrocyte sedimentation rate (ESR) and patient's perceived general health [11].

The precise timing of when to initiate treatment, a so called "Window to treat", is not well defined and has changed over the years [12]. The generally accepted goal is to prevent irreversible damage to the joints and body, and treatment initiated as early as possible is of vital importance [13, 14]. In order to treat RA as early as possible, an early diagnosis is necessary, which in itself poses challenges due to that the non-specificity of early disease signs and symptoms of RA can be overlapping with other diseases such as inflammatory arthritis, psoriatic arthritis and osteoarthritis [7], which makes it challenging to discriminate between RA, other inflammatory as well as non-inflammatory joint disorders.

Furthermore, there is evidence to suggest that the disease process starts years before the first

symptoms are manifest. The presence of autoantibodies to citrullinated proteins (ACPAs) and antibodies to immunoglobulins (rheumatic factor; RF) are closely associated with RA and can be detected decades before any clinical symptoms [15, 16]. Together with typical disease symptoms, these autoantibodies helps the physician to diagnose the patients correctly. However, approx. 30% patients are “seronegative”, which means they do not have either of these antibodies [17] that sometimes complicate the ability to correctly classify the patients.

Classification criteria

As there are no tests or biomarkers that can with certainty determine if a patient has RA, the diagnosis is often established based on classification criteria, which are updated at regular intervals, and have thus changed over the years. Nevertheless, classification criteria are not the same as diagnostic criteria. Unlike diagnostic criteria, the goal of the classification criteria is to select a homogenous group of patients that can be studied from a research perspective and do not necessarily have to accurately identify all patients with the disease. They are therefore less broad and more focused on capturing key features of the condition [18].

One such set of classification criteria was developed in 1987 by the American College of Rheumatology (ACR) [19] and includes the following features: 1. Morning stiffness 2. Arthritis of three or more joint areas 3. Arthritis of hand joints 4. Symmetric arthritis 5. Rheumatoid nodules 6. Rheumatoid factor 7. Radiographic changes. If a patient fulfils at least 4 out of these 7 criteria during at least 6 weeks, she or he is considered to have RA. A problem with the 1987 ACR criteria is that they define a patient with established disease but are not so useful for early identification. Therefore, the European League Against Rheumatism (EULAR), together with the ACR, developed a new set of criteria in 2010 that

tries to remedy this shortcoming [20]. These criteria are focused on earlier signs of inflammation and arthritis, and high titer of anti-citrullinated protein antibodies (ACPA) and immunoglobulin G (rheumatic factor; RF) have a considerable weight (Table 1). Eligible patients are those with a swollen joint that cannot be explained by another disease. Patients with typical radiological damages that suggest RA are classified as having RA irrespective the algorithm

Table 1. Algorithm for ACR/EULAR 2010 classification criteria for RA. The patient is classified as having RA when the score is or exceeds 6.

Joint involvement*	
1 large** joint	0
2-10 large joints	1
1-3 small*** joints (with or without involvement of large joints)	2
4-10 small joints (with or without involvement of large joints)	3
>10 joints (at least one small joint)	5
Serology	
Negative ACPA and RF	0
Low positive ACPA and/or low positive RF	2
High**** positive ACPA and/or high positive RF	3
Acute phase reactants	
Normal CRP and ESR	0
Abnormal CRP or ESR	1
Duration symptom***	
< 6 weeks	0
≥ 6 weeks	1

*Joint involvement refers to any swollen or tender joint on examination, which may be confirmed by imaging evidence of synovitis. Distal interphalangeal joints, first carpometacarpal joints, and first metatarsophalangeal joints are excluded from assessment.
**Large joints" refers to shoulders, elbows, hips, knees, and ankles.
***Small joints" refers to the metacarpophalangeal joints, proximal interphalangeal joints, second through fifth metatarsophalangeal joints, thumb interphalangeal joints, and wrists.
****>3 times the ULN for the laboratory and assay.
*****Duration of symptoms refers to patient self-report of the duration of signs or symptoms of synovitis (e.g., pain, swelling, tenderness) of joints that are clinically involved at the time of assessment, regardless of treatment status.
Adapted from https://www.rheumatology.org/Practice-Quality/Clinical-Support/Criteria assessed 220406

There are no classification criteria that have 100% sensitivity and specificity, and some patients are wrongly diagnosed with RA. On the one hand, the EULAR/ACR 2010 criteria are beneficial as they allow early identification of patients and early treatment can be started but on the other, the likelihood of a wrong diagnosis has increased. In one study as much as 10% of patients initially diagnosed with RA were given a definite alternative diagnosis [21].

Misdiagnosis

In this work, we examine the misdiagnosis in a large cohort of patients diagnosed with early RA using the ACR 1987 criteria. A cohort study where patients are followed over time (sometimes over many years) with repeated assessments provides a good framework for studying the extent of misclassification in RA. In this project, we have used data from the BARFOT (Better Anti Rheumatic Pharmaco Therapy) cohort, a multicentre long-term study that included more than 2800 early RA patients from six clinical centres in Sweden who have been followed for up to 15 years [22]. A number of these patients had a change in diagnosis to other inflammatory diseases at a later time point. We aimed to characterize these patients in detail in order to evaluate whether there are clinical factors that can help facilitate a correct diagnosis and prevent wrong diagnosis in the future.

Among the risk factors for misclassification that we considered were each of the ACR 1987 classification criteria [19], relevant clinical variables such as age at diagnosis, gender, smoking, number of swollen and tender joints, as well as presence of serum RF [23], presence of anti-cyclic citrullinated peptide (anti-CCP)-antibodies, i.e. ACPA [24] and DAS28 [25].

Specific Research questions

Based on the follow-up data we were able to identify patients who were correctly diagnosed with RA (hereafter called RA-keep) and those who were misdiagnosed (hereafter called RA-change).

We examine differences between the groups in regard to:
Disease and patient characteristics at diagnosis; number and type of ACR1987 criteria fulfilled; which joints were first affected; presence of ACPAs; disease outcome during the follow-up time determined as disease activity and radiological damage; choice of anti-rheumatic treatment. We also compared the time in the study from diagnosis in the two groups.

Material and Methods

Description of the cohort

All subjects in this study were taken from a Swedish early RA cohort, BARFOT (Better Anti Rheumatic FarmacO Therapy) who fulfilled at least four of the seven ACR 1987 criteria and had a symptom duration $\leq$ 12 months. Follow-up time was scheduled to 15 years. Approximately half of the patients were enrolled during 1990's and the other half during the 2000's, n=2838 in total (Table 2). The differences between the patients included during the two periods were limited [26] and to gain enough power, all eligible patients were included in the present study. The inclusion criteria for the present study were patients in BARFOT who fulfilled at least four ACR1987 classification criteria, had at least four clinical visits and a symptom duration of less than 12 months were included in our study, n=2541. RA was classified according to the ACR 1987 criteria: 1) Morning stiffness; 2) Arthritis of 3 or more joint areas; 3) Arthritis of

hand joints; 4) Symmetric arthritis; 5) Rheumatoid nodules: g) Serum RF and 7) Radiographic erosion. When at least 4 out of these 7 criteria were fulfilled and the symptoms had been ongoing for at least 6 weeks, the patient was classified as having RA.

The individual ACR1987 RA classification criteria were assessed during the 15-year follow up period. Due to power issues in the RA-change group, the following clinical parameters were analysed at onset, at 3 and 6 months, and at 1, 1.5 and 2 years: DAS28 is a composite score that asesses the number of swollen and tender joints out of 28 joints, ESR and patient's own assessment of global health on a visual analog scale (VAS, 0-100 representing best to worst). In addition, gender, age at inclusion and smoking, and use of disease modifying anti-rheumatic drugs (DMARDs) and glucocorticosteroids (GC), as well as perceived VAS pain (0-100, best to worst).

Joints were also assessed as to first affected joints. The joint or joints that were first affected were categorised into three different groups: Small joints (joints in hands, wrists and feet), Large joints (elbows, shoulders, hips and knees) and other/unknown (joints in neck, extra articular and tendinitis). The distribution of first affected joints was compared between the RA-keep and RA-change groups (Table 3). Of the 2541 patients in the study, data on affected joints was missing for 432 patients.

Radiographs of hands and feet were taken at onset and after 1 and 2 years. The radiology joint destruction was scored according to modified Sharp van der Heijde Score (mSHS) by a

blinded qualified assessor. mSHS is a composite score of joint narrowing score (JNS, i.e., cartilage destruction) and bone erosion (ES) that together result in the total radiologic score.

Antibodies to citrullinated proteins antigens (ACPA) measured as anti-CCP was not a routine parameter during the collection of the cohort and were analyzed from biobanked serum at later time points either by an ELISA from Euro-Diagnostica, Malmö, Sweden or by a routine assay at Clinical Immunology at Sahlgrenska University, Göteborg, Sweden.

Table 2: Characteristics at inclusion of the whole BARFOT cohort. Statistics are presented as number (percentages) for categorical variables (gender, smoking, RF, Anti-CCP, treatments) and median with percentile (P_{25}-P_{75}) for continuous variables (Age, DAS28, Global Health, Swollen joint count, ESR).

Parameter	Frequency/ Median
Total number of participants, n=2838	
Female	1916 (67.5%)
Male	922 (32.5%)
Smoking status, n=2735	
Ever smoking	1644 (60.1%)
Never smoking	1091 (39.9%)
Age at first symptom	60 (48 - 71)
Symptom duration	6 (4-9)
DAS28	5.3 (4.5 - 6.1)
Global Health (VAS)	46 (25 - 64)
Swollen joint count	10 (6 - 14)
Tender joint count	7 (3 - 12)
ESR	30 (16 - 50)
RF status, n=2779	
Positive	1682 (60.5%)
Negative	1097 (39.5%)
Anti-CCP status, n =1936	
Positive	1133 (58.5%)
Negative	803 (41.5%)
Treatment, n=2838	
Methotrexate	1239 (43.7%)
Other synthetic	888 (31.3%)
Glucocorticosteroids only	285 (10%)
Combination	53 (1.8%)
Biologic DMARDs	12 (0.4%)
None	360 (12.8%)

DAS28 - Disease Activity Score, VAS – visual analogue scale, ESR - erythrocyte sedimentation rate, RF - rheumatoid factor, Anti-CCP - Anti-cyclic citrullinated peptide, DMARDs - Disease-modifying antirheumatic drugs.

Criteria for identification the RA-keep and RA-change groups

According to our inclusion criteria for the present study, we could identify two groups of patients: those who kept their RA diagnosis throughout the follow-up (RA-keep) and those whose diagnosis changed (RA-change).

Identification of the RA-keep group: Patients in the BARFOT cohort were defined as the RA-keep group based on the following inclusion criteria: fulfilling ≥4 of the ACR1987 criteria, have at least 4 clinical visits – where non-missing visit means having information on DAS28, pain or HAQ and who did not go into remission before 6 months. A total of 2497 patients were defined as RA-keep during the 15-year follow-up time.

Identification of the RA-change group: Patients in the BARFOT cohort were defined as the RA-change group based on the following inclusion criteria: patients fulfilling ≥4 of the ACR1987 criteria, have at least 4 clinical visits – where non-missing visit means having information on DAS28, pain or HAQ and had a changed diagnose or went into remission before 6 months. A total of 44 patients were defined as RA-change during the 15-year follow-up time.

The total number of patients to be included in this study is 2541. From the 2838 patients from the BARFOT cohort a total of 297 were excluded for not fulfilling the inclusion criteria.

Statistical methods

Some of the factors of interest, that we chose according to the specific research questions, were continuous and some were categorical. These were compared between the groups RA-keep and RA-change. Appropriate statistical tests were used depending on the type of parameter. Thus, we have compared the medians in the two groups for all the continuous parameters, such as age at first symptom, DAS28 score, number of swollen or tender joints, by using the Mann-Whitney U-test. Because the two groups have very different sizes (2497 vs 44), we chose to compare the medians and not the means, since the mean in the smallest group could be much more affected by the outliers compared to the mean in the largest group.

For categorical parameters, such as smoking, gender, presence of anti-CCP and fulfillment of classification criteria, differences in proportions were compared with the Chi-square test. For those instances where one off the groups in the comparison included 5 or less participants (ex. RA-change and anti-CCP negative group) we used Fisher's exact test to obtain a more accurate p-value.

After we identified the parameters that were significantly different between the two groups, we assessed the size of the effect that the respective parameter (example, RF positivity) had on the probability of being in the RA-keep group. We have done this by using logistic regression models, where we have specified the RA-keep group as the outcome of interest, the RA-change group as the reference category for the outcome of interest and RF positivity as the explanatory variable in the regression model. The result of such a model is the odds ratio (OR), together with its 95% confidence interval (CI), for being in the RA-keep group for the

patients that were RF positive. The OR estimates in this case how many times more likely (or less likely when OR<1) it is to be in the RA-keep compared to being in the RA-change group. From the OR we could also calculate the probability to belong to the RA-keep group according to formula (1):

$$P(\text{RA-keep}) = OR/(1+OR) \qquad (1)$$

where OR= OR (being in the RA-keep) and P (RA-keep) =probability of being in the RA-keep group.

Linear repeated measure models

In order to assess differences over time (0-24 months) between the two groups regarding the clinical outcomes of interest such as DAS28 score, number of swollen and tender joints, ESR, global health and VAS pain we used linear repeated measure models for each such outcome. Repeated measure models are a suitable way to analyze measurements that are collected at more than one time point for the participants, as the measurements of each participant are correlated between themselves.

In these analyses, the mean evolution of for example DAS28 is estimated separately for the RA-keep and RA-change groups as a linear trend. The difference between these two trends is assumed constant over time (more advanced models are possible when the difference is allowed to vary over time). A significant p-value from this model means that the evolution of the two groups over time is significantly different.

Furthermore, two types of linear repeated models were employed. In the first type, all measurements (0-24 months) were used to answer the question whether there is a significant difference between the evolution of the two groups over time. If such a difference was found,

a second model was employed where only the 3-24 months measurements were included, and the model was adjusted instead for the inclusion measurement. This model addresses the question: given that there is a significant difference in the evolution over time, could this difference be explained by a difference present already at inclusion? If the p-value for the groups is not significant it means that the difference in evolution is explained by the difference at inclusion (i.e. the groups continue with the same tendency they that had at inclusion). If it is significant it means that the evolution of the two groups is significantly different and the difference is bigger (or smaller) than could have been expected from the inclusion.

Ethics

This is an observational study and participation did not affect the treatment or care for the patients. All patients signed informed consent forms, people not able to give informed consent were not included (Ethical permit number 1994-11-16, LU368-94, Gbg 88-94, Li 94283). To mitigate the risk of privacy violation all data is stored in password-protected computers with strictly restricted access. All personal data is pseudonymized and is handled in accordance with the Data Protection Act.

Results

Characterisation of RA-keep and RA change

The characteristics that differed significantly or showed a different tendency between the RA-keep and RA-change groups are showed in Table 3. The groups have a similar distribution regarding gender and age at first symptom, but the RA-keep group has more patients that had ever smoked and had a higher median DAS28 at inclusion compared to the RA-change group (p=0.001). Tender joint count and ESR were higher in RA-keep than RA-change.

Table 3: Characteristics at inclusion of the RA-keep and RA-change groups. Statistics are presented as number (percentage) for categorical variables (gender, smoking) and median with percentile (P25- P75) for continuous variables (age at first symptom, DAS28 at inclusion).

	RA-keep (N=2497)	RA-change (N=44)	p-value
Ever smoker N, (%)	1464 (60%)	17 (42%)	0.02
Age at first symptom	58 (47 - 69)	64 (46 - 72)	0.32
Symptom duration	6 (4 - 9)	4.5 (3 - 6)	<0.01
DAS28 at inclusion	5.3 (4.5 - 6.1)	4.7 (3.8 - 5.5)	0.001
Tender joint count	7 (3 - 12)	4.5 (2 - 12)	0.10
ESR	30 (17 - 50)	19.5 (8.5 - 32)	<0.001

DAS28 - Disease Activity Score, ESR – Erythrocyte Sedimentation Rate

Time in the study for the two groups

The amount of time in the study from inclusion until date of exclusion or 31st December 2019, whichever came first, was investigated for the two groups with the help of Kaplan-Meier survival analysis (Fig. 1). Time in the study for the RA-change group was significantly shorter compared to the RA-keep ($p<0.001$). Nevertheless, 5 years after inclusion more than 50% of the RA-keep group was still in the study and more than 20% of the RA-change patients are still considered RA 10 years from inclusion.

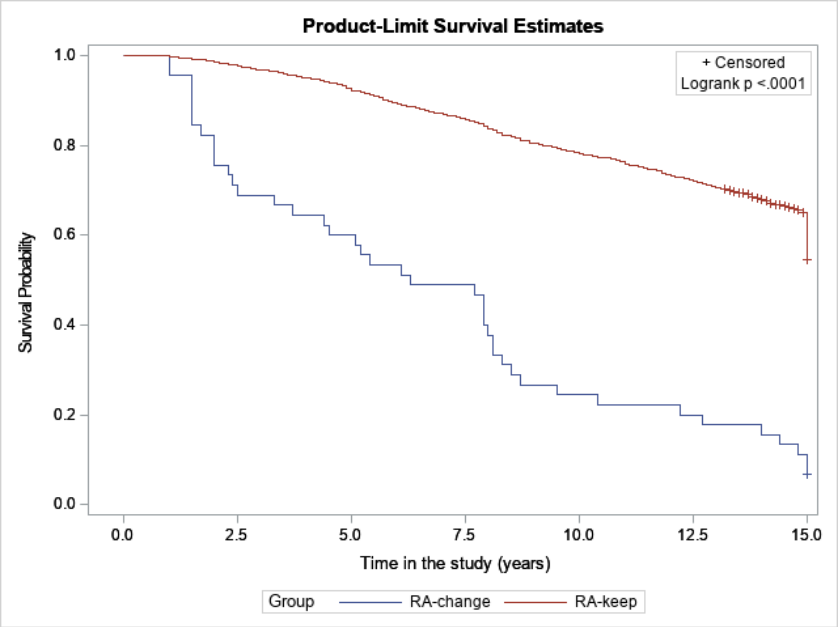

Figure 1: Kaplan-Meier survival curves for the RA-keep and RA-change groups. Time represents the time in the study. P-value <0.001

Differences in fulfillment of ACR 1987 classification criteria between RA-keep and RA-change

As mentioned before, RA was classified according to ACR1987 criteria. We analyzed whether there was any difference regarding which specific criteria that the RA-keep and RA-change group fulfilled. Of the seven individual criteria the two criteria for which the two groups differed were the RF-positivity (RA-keep 63.1%, RA-change 21.4%, p=<0.001) and radiographic changes at inclusion (RA-keep 26.5%, RA-change 12.2%, p=0.04).

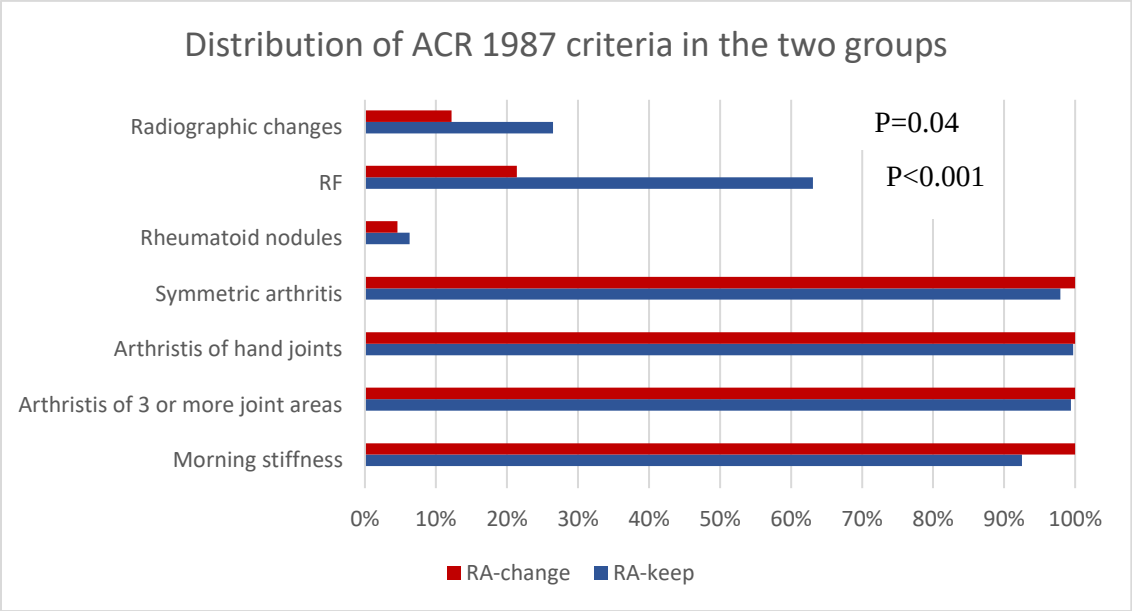

Fig. 2 Shows the distribution of fulfilled criteria for the two groups, RA-change and RA-keep.

Percentage of patients positive for RF is showed in Fig 3. According to the univariate logistic regression model, the OR for being in the RA-keep group when being positive for RF was 6.3 (95% CI 2.98-13.14), p<0.001, which means that being RF positive increases 6 times the odds to have a true RA disease. Further, the probability for a RF positive patient to belong to the RA-keep group was 86%.

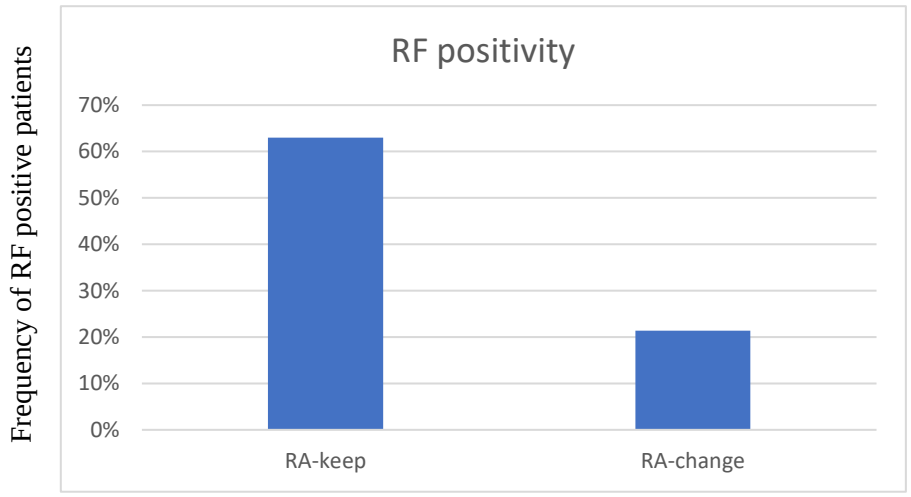

Figure 3: Percentage of RF positivity in RA-keep (63 %) and RA-change (21.4%). P-value <0.001

To assess whether the number of fulfilled ACR1987 criteria differed between the two groups, we have furthered grouped patients into those who fulfilled 4 criteria and those who fulfilled more than 4 criteria. The number of patients that fulfilled more than 4 ACR1987 criteria was significantly higher in the RA-keep group compared to the RA-change group (RA-keep 63.5%, RA-change 34.1%, p=<0.001) (Fig. 4).

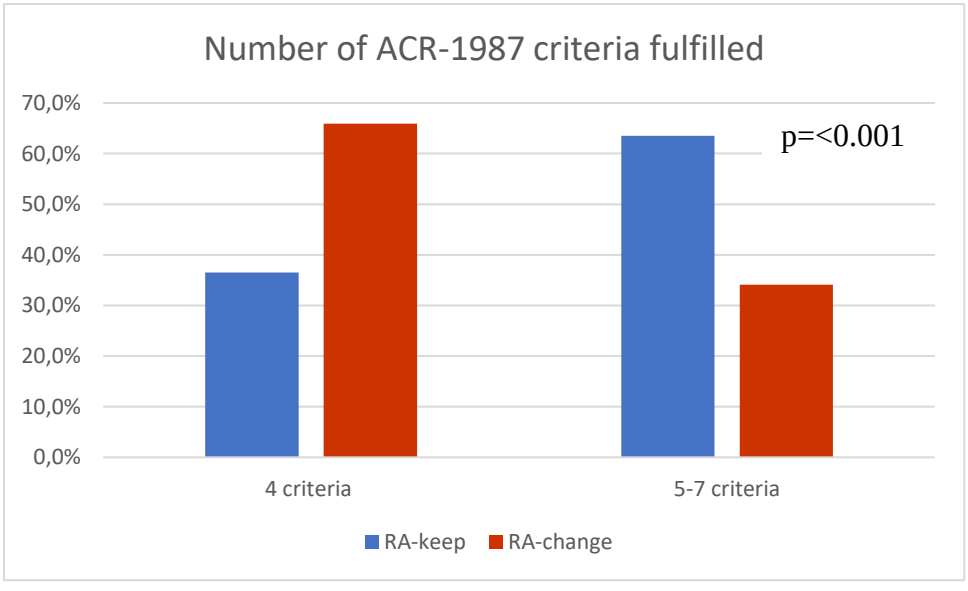

Figure 4: Distribution of number of fulfilled criteria by group. P-value <0.001

Presence of anti-CCP antibodies in RA-keep and RA-change

Despite a significant amount of missing data regarding anti-CCP status (30% of the patients), it is noteworthy that there was a significantly higher percentage of anti-CCP positive in the RA-keep group (60 %) than in RA-change (9 %) ($p<0.001$).

Difference in distribution of first affected joints between RA-keep and RA-change

First affected joints were categorised into small joints, large joints and other or unknown as described in the method section and compared the RA-change and RA-keep groups.

When calculated with Chi-square we saw no differences between RA-change and RA-keep regarding first affected joints when comparing small and large joints ($p=0.51$). The group for other/unknown for RA-change was too small for the calculation (Table 4).

Table 4: Number and percentage of the first affected joint-group in RA-Keep and RA-Change. Total number of patients with available information n=2109. Percentages are presented within RA-keep and RA-change, respectively.

	RA-Keep	RA-Change	p-value*
Small joints n, (%)	1454 (70%)	23 (68 %)	0.51
Large joints n, (%)	546 (26%)	11 (32 %)	ns
Other/ unknown n, (%)	75 (4 %)	0 (0 %)	NA
Missing information, n=432			

**P-value was given according to the Chi-square test of association. Category 'Other/ unknown' was not included in the test*

Differences in disease activity between patients in the RA-keep and the RA-change groups

We examined whether the evolution of the composite measure of disease activity DAS28 (Fig. 5A) and its individual components that includes patient's global health (Fig. 5B), the number of swollen joints (Fig. 5C), the number of tender joints (Fig. 5D) and ESR (Fig. 5E) differed over between the two groups of interest during the first two years after inclusion in BARFOT. Differences in the evolution of these parameters were estimated by linear repeated measure models. Of the assessed parameters, ESR evolution was the only one that differed significantly between the groups ($p=0.02$), with the RA-keep being 5.5 units higher than group RA-change group over time. However, when only the 3-24 months evolution was analyzed and the model was adjusted for ESR at inclusion, the effect of the group was no longer significant suggesting that the difference between the groups could be explained by baseline differences.

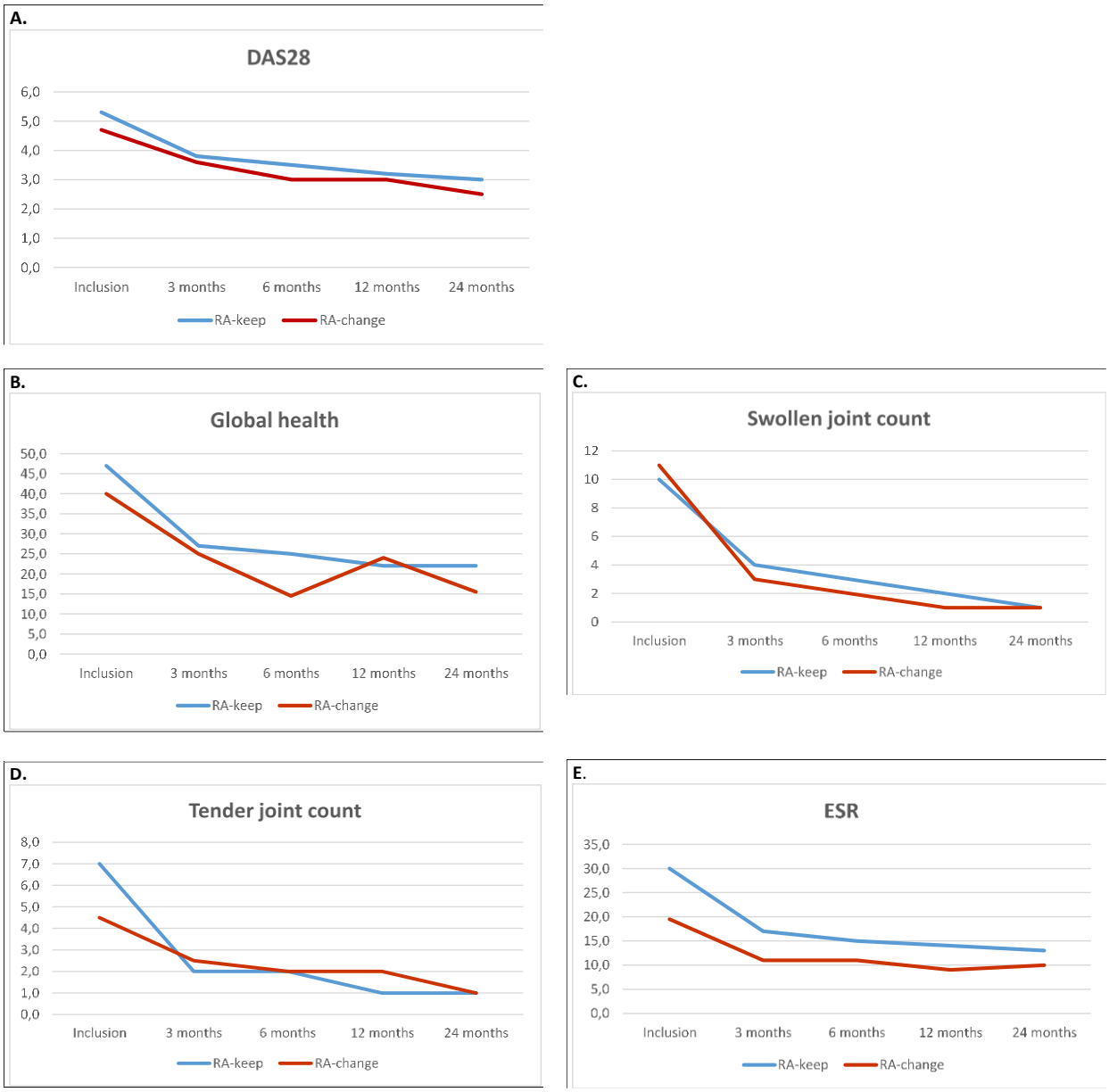

Figure 5: Median value for DAS28 together with clinical components over two years of follow-up of the RA-change and RA-keep groups. A. The total DAS28 score, B. The number of swollen joints out of 28 assessed, C. The number of tender joints out of 28 assessed, E. ESR. DAS28 – Disease activity score 28, ESR – erythrocyte sedimentation rate

Regarding perceived VAS pain over time there were no significant differences between the

RA-keep and RA-change groups (Fig. 6).

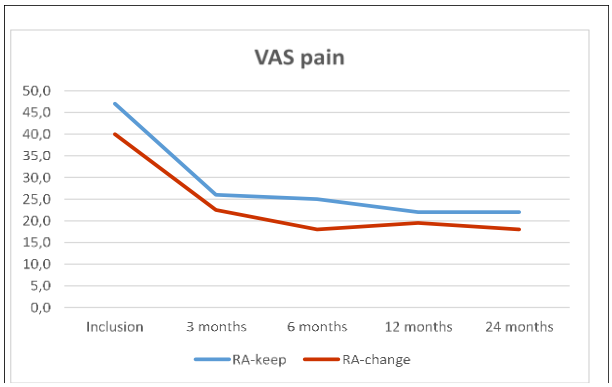

Figure 6. Median value for pain measured as VAS over two years of follow-up by group (RA-change and RA-keep). VAS – visual analogue scale. Linear repeated measures model

Differences in joint destruction between the RA-keep and the RA-change group

There were no differences in the proportion of patients between the groups with respect to JNS and the total mSHS at inclusion. However, the proportion of patients who presented with erosion (ES) differed between the groups: In the the RA-keep group 28% had erosions at inclusion compared to 9% in the RA-change group (p=0.04). Evolution of radiological damage over time was limited to 24 months. Both the total radiological score (Fig. 7A) and the JNS (Fig. 7B) increased more in the RA-keep than in the RA-change group although the results did not reach the level of significance, probably due to lack of power in the RA-change group.

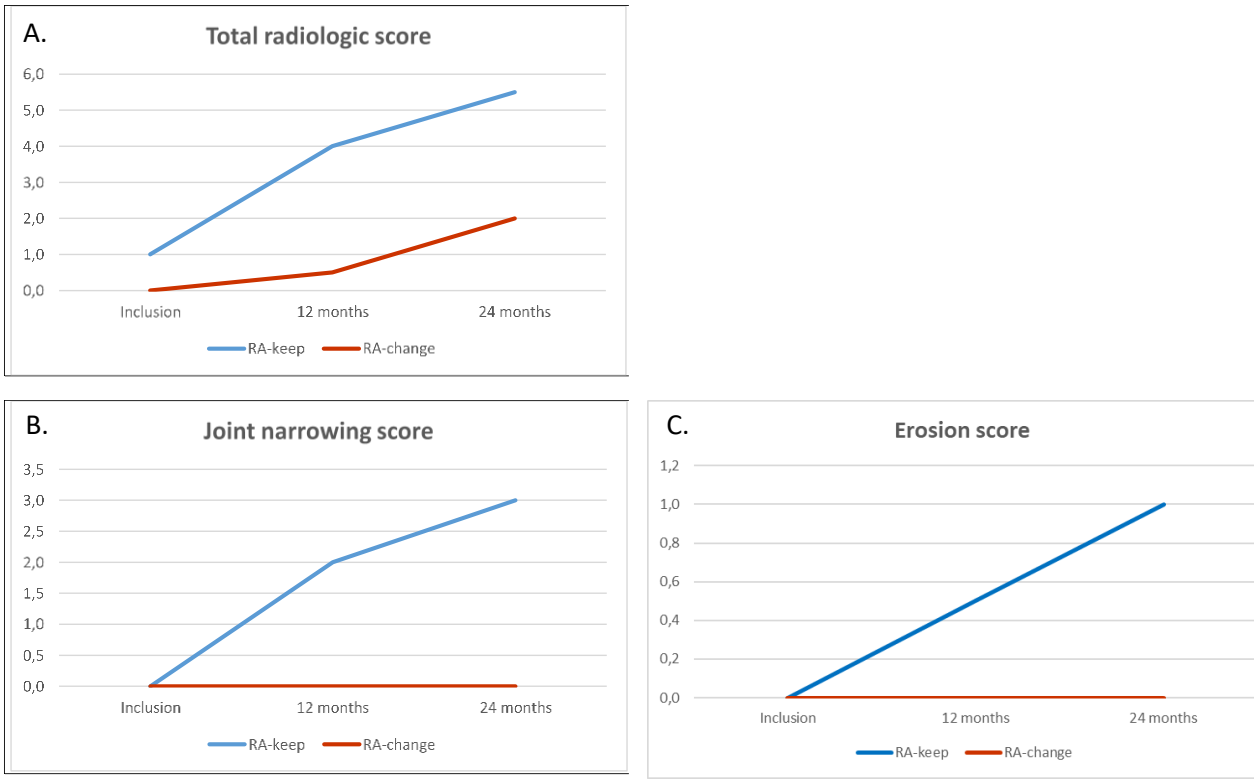

Figure 7: Median value for radiologic destruction in the RA-keep and the RA-change groups. A. Total radiological score (mSHS), B. Joint narrowing score (JNS),C. Erosion score (ES) over two years of follow-up. Repeated measure model

Differences in the treatments given to the RA-keep and RA-change groups at inclusion

We studied whether there were any differences in the proportion of patients in the two groups who received DMARDs at inclusion (Table 5). A significant lower proportion of the patients in the RA-keep group received any DMARDs compared to those in the RA-change group (21% vs 43%, $p<0.01$). In the RA-keep group 46% of patients received methotrexate compared to 32% in the RA-change group.

Table 5. Treatment given at inclusion for the RA-Keep and RA-Change groups. Statistics are given as number and percentage (%) within the respective group.

Medication	**RA-Keep**	**RA-Change**	**p-value***
None	520 (21%)	19 (43%)	<0.01
Methotrexate	1121 (46%)	14 (32%)	
Other synthetic	797 (33%)	11 (25%)	
Total	2438 (100%)	44 (100%)	

Note: 59 patients in the RA-Keep group who received Biologicals or combination therapy were not included in this comparison due to a lack of a similar group within the RA-change group.

Discussion

Misdiagnosis of RA

Diagnosis of RA is problematic, and a wrong diagnosis might lead both to over and under treatment. In this study we found the only 1.7% of the patients in the BARFOT study changed their RA diagnosis during the follow-up time. However, our main findings are that seropositivity, i.e., the presence of RF together with fulfilment over more than 4 ACR 1987 classification criteria increases the likelihood of a correct diagnosis from inclusion.

The treatment of autoimmune inflammatory arthritidis is similar, but during the last decade the diseases differences have become more and more clear, which can be exemplified by the

fact that e.g., IL17 inhibitors work very well in spondylarthritis such as psoriatic arthritis but has a very limited effect on RA. Thus, it becomes more and more clear that the different types of inflammatory arthritidis actually are very different diseases entities with different disease pathologies. This means that the demands for a correct diagnosis have increased, and misdiagnosis might lead to reduced quality of life due to not addressing the real underlying disease, together with high costs for the society in providing social and medical support.

The proportion of patients in BARFOT who changed their diagnosis over the study period was very small (1.7%). This is probably due to the fact that BARFOT was a clinical study with defined inclusion criteria (RA according to ACR 1987 criteria). In contrast, in one cross sectional study including over 4000 patients with a presumptive RA diagnosis who were referred from primary care to a specialized RA center for confirming the diagnosis, ss much as 39% of these were actually misdiagnosed, with the most prevalent diseases in this group being osteoarthritis (50%), systemic lupus erythematosus (5%) and Sjögren's syndrome (3%) [27]. Of note is also that infections, such as *Thropheryma wipplei* bacteria that cause Whipple's disease [27], Chikungungya [28] and Parvovirus B19 [29] can mimic RA, and suggested that especially seronegative patients with a poor treatment response should be reevaluated. Identifying possible differences between the groups RA-keep and RA-change could contribute to identifying risk factors for misdiagnosis.

In this study, that includes more than 2500 patients with early diagnosed RA we investigated if there are differences in characteristics between the patients correctly diagnosed with RA (the RA-keep group) and those who have other inflammatory diseases and were misdiagnosed as RA (the RA-change group). The patients were included in the BARFOT cohort if they

fulfilled the ACR1987 classification criteria and had a symptom duration of less than 12 months.

Further, we assessed which of the disease and patient characteristics at diagnosis could be helpful in setting a correct diagnosis. These characteristics included the number of ACR1987 criteria fulfilled, which of the ACR 1987 criteria were fulfilled, the first affected joints. Other tested characteristics were presence of ACPAs, choice of anti-rheumatic treatment, as well as disease activity (DAS28) together with its components and radiological destruction both at diagnosis and at follow-up times.

Number and type of ACR1987 criteria fulfilled

We assessed whether the number of fulfilled ACR 1987 criteria differed between RA-keep and RA-change and found that the RA-keep group was much more likely to fulfilled more than 4 of the criteria (63.5%) whereas in the RA-change group a much lower proportion (34.1%) fulfilled more than 4 of the criteria (Figure 4). These findings are not surprising, RA is a disease that develops over time with more and more symptoms. Indeed, the RA-keep group had a longer symptom duration before inclusion compared to the RA-change group and had develop a more obvious disease. The criteria that differed significantly between the groups were RF positivity and radiographic changes. These findings are interesting as the newer EULAR/ACR 2010 [20] criteria immediately classify a patient with typical joint destruction as RA and are also heavily based on the presence of autoantibodies (RF and ACPA). In addition, both RF and ACPA are associate with a more destructive disease, thus these two criteria might be linked.

In general, approx.. 70% of the patients with RA are RF and/or ACPA positive. In our comparison between the RA-keep and RA-change groups the presence of RF was significantly higher in the RA-keep group (Fig. 3). Although RF is found in 5% of the healthy population and can be present in patients with other autoimmune diseases such as systemic lupus erythematosus and systemic sclerosis as well as in chronic infections and in non-autoimmune diseases [23], it has a sensitivity of 60-90% and specificity 48-92% for RA, and facilitates a correct diagnosis [5] . RF can be present in both IgM, IgG and IgA isotypes, IgM is the one that is most clinically used RF is that it is found in multiple isotypes (IgM, IgG and IgA), with IgM being the most commonly measured in clinical settings. More than half of the RA patients present all three isotypes, whereas less than 5% of the healthy individuals positive for any RF are positive for all three isotypes [22].

ACPA, i.e. antibodies to citrullinated protein antibodies, have been identified as an important predictive factor for correct diagnosis of RA patients, with a sensitivity of 67% and specificity 95% for diagnosis [30]. Patients who are negative for both ACPA and RF are diagnosed later in the disease course and have a higher degree of inflammation at diagnosis [31]. One study investigated over 4000 patients from two cohorts (ESPOIR and Leiden-EAC) who were diagnosed with RA or undifferentiated arthritis (UA) according to the ACR1987 criteria [32]. Of the UA patients, approximately 25% (n=463 patients) could be classified as RA one year later. However, when using the EULAR2010 criteria 75% of these could have been classified as RA already at inclusion and those who were ACPA-positive had much higher chance of being correctly identified earlier. The study confirms that RA can be identified earlier with the EULAR 2010 criteria compared to the ACR1987 criteria and adds the information that the

2010 criteria work especially well in identifying RA in autoantibody positive patients but not so well in autoantibody negative patients [33]. These studies are in line with our results that indicate that 60% of the patients in the RA-keep group were ACPA positive at the same time as the proportion of ACPA positive patients in the RA-change group was 9%. However, this marker was not being routinely measured in the clinic before 2006 and all patients in the BARFOT cohort were already diagnosed before that time point. ACPA analysis was only available for approx. 70% of the patients. Due to the low number of participants in the RA-change group, together with the risk that the missing data could be non-random (i.e., higher probability that ACPA would be less assessed in the RA-change group) made us cautious in putting weight on this finding.

First affected joints

We investigated which joints that were first affected but found no difference between the groups. Small joints are affected is also seen in osteoarthritis and in other inflammatory diseases such as reactive arthritis, systemic lupus erythematosus and psoriasis arthritis although with a different distribution pattern [1]. The BARFOT database does not allow discrimination between the different finger joints, and together with coupled with overlapping inflammatory symptoms, these analyses could not discriminate between the RA-keep group and the RA-change group.

Disease outcome according to DAS28 during the follow-up time

There was a significant difference between the groups in DAS28 at inclusion, with a higher score for RA-keep than RA-change ($p=0.001$). Of the individual DAS28 components, ESR

was significantly higher for RA-keep at inclusion (p=<0.0001), indicating a higher disease activity in the RA-keep group than in the RA-change. As for the follow-up of 24 months none of the components showed a significant difference. The patients in the RA-change group had their RA diagnosis changed to different other diseases even though their symptoms were initially similar to RA, and this could be the reason for some of them to score lower in DAS28.

Choice of anti-rheumatic treatment

A higher proportion in the RA-change group received no medication (except glucocorticosteroids) at inclusion and a lower proportion received methotrexate which is the recommended treatment to initiate immediately after diagnosis of RA throughout most of the BARFOT study. Initiating medical treatment within the first months after diagnosis of RA is essential for a good outcome of the "treat to target" strategy where a low disease activity or even remission is the aim [30]. We hypothesized that the difference in medication between the groups could be explained by the time of diagnosis. We analyzed if there was a difference in time for diagnosis between the groups, i.e., whether the patients had been included in BARFOT during the 1990's or during the 2000' but no major difference were found [26]. The higher score in DAS28 for the RA-keep group implied that these patients either had a more manifest disease (more joint destruction and longer disease duration), which could explain the higher proportion of medical treatment at inclusion in this group.

Strengths and limitations

The major strength of this study is the very well characterized and meticulously collected BARFOT cohort with very little missingness and a follow-up period of 15 years. Not many other longitudinal studies have such an extensive number of patients who are followed over such a long period of time. Although one limitation is the low number of participants in the RA-change group, smaller cohort than BARFOT would not have allowed a study such as ours. Unfortunately, ACPA was not available for all patients. Despite an interesting lower trend of radiologic outcome in the RA-change compared to RA-keep group, we had to limit these analyses at 2 years due to power issues. The small number of participants in the RA-change group did probably result in a lack of statistical significance regarding the radiologic differences between the groups.

Conclusions

Diagnosis early in the disease, before the symptoms become clearer, is difficult. The first symptoms can be very subtle, diffuse and vary greatly from person to person. They also overlap or resemble symptoms of many other inflammatory or autoimmune diseases which makes diagnosis difficult. Our primary findings indicate that for patients that are (both) RF negative and do not fulfil more than 4 of the ACR1987 criteria, diagnosis of RA should probably be reconsidered.

Our study was limited by the small number of misdiagnosed patients in the RA-change group. If a similar study was conducted according to the newer EULAR/ACR2010 criteria, a larger

proportion of misdiagnosed might be found due to the lower specificity in this set of criteria. This could perhaps show differences with stronger significance. We believe that studies like ours could be of help for clinicians to avoid misdiagnosis that creates unnecessary mental stress for the patient, leads to wrong treatment and continued suffering as the real disease is not addressed.

References

1. Sparks, J.A., *Rheumatoid Arthritis.* Ann Intern Med, 2019. **170**(1): p. ITC1-ITC16.
2. Cross, M., et al., *The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study.* Ann Rheum Dis, 2014. **73**(7): p. 1316-22.
3. Oton, T. and L. Carmona, *The epidemiology of established rheumatoid arthritis.* Best Pract Res Clin Rheumatol, 2019. **33**(5): p. 101477.
4. Smolen, J.S., et al., *Rheumatoid arthritis.* Nat Rev Dis Primers, 2018. **4**: p. 18001.
5. Pollard, K.M., *Silica, Silicosis, and Autoimmunity.* Front Immunol, 2016. **7**: p. 97.
6. Henry, J., E. Roulot, and C. Gaujoux-Viala, *[The rheumatoid hand].* Presse Med, 2013. **42**(12): p. 1607-15.
7. Smolen, J.S., D. Aletaha, and I.B. McInnes, *Rheumatoid arthritis.* Lancet, 2016. **388**(10055): p. 2023-2038.
8. Wasserman, A.M., *Diagnosis and management of rheumatoid arthritis.* Am Fam Physician, 2011. **84**(11): p. 1245-52.
9. Matcham, F., et al., *The impact of rheumatoid arthritis on quality-of-life assessed using the SF-36: a systematic review and meta-analysis.* Semin Arthritis Rheum, 2014. **44**(2): p. 123-30.
10. Yu, C., et al., *Remission rate and predictors of remission in patients with rheumatoid arthritis under treat-to-target strategy in real-world studies: a systematic review and meta-analysis.* Clin Rheumatol, 2019. **38**(3): p. 727-738.
11. van Riel, P.L. and L. Renskers, *The Disease Activity Score (DAS) and the Disease Activity Score using 28 joint counts (DAS28) in the management of rheumatoid arthritis.* Clin Exp Rheumatol, 2016. **34**(5 Suppl 101): p. S40-S44.
12. Burgers, L.E., K. Raza, and A.H. van der Helm-van Mil, *Window of opportunity in rheumatoid arthritis - definitions and supporting evidence: from old to new perspectives.* RMD Open, 2019. **5**(1): p. e000870.
13. Nell, V.P., et al., *Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis.* Rheumatology (Oxford), 2004. **43**(7): p. 906-14.
14. van der Heide, A., et al., *The effectiveness of early treatment with "second-line" antirheumatic drugs. A randomized, controlled trial.* Ann Intern Med, 1996. **124**(8): p. 699-707.
15. Aho, K., et al., *When does rheumatoid disease start?* Arthritis Rheum, 1985. **28**(5): p. 485-9.
16. Rantapaa-Dahlqvist, S., et al., *Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis.* Arthritis Rheum, 2003. **48**(10): p. 2741-9.
17. De Winter, L.M., et al., *Autoantibodies to two novel peptides in seronegative and early rheumatoid arthritis.* Rheumatology (Oxford), 2016. **55**(8): p. 1431-6.
18. Aggarwal, R., et al., *Distinctions between diagnostic and classification criteria?* Arthritis Care Res (Hoboken), 2015. **67**(7): p. 891-7.

19. Arnett, F.C., et al., *The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis.* Arthritis Rheum, 1988. **31**(3): p. 315-24.
20. Aletaha, D., et al., *2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative.* Arthritis Rheum, 2010. **62**(9): p. 2569-81.
21. Berglin, E. and S.R. Dahlqvist, *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): p. 362-8.
22. Hafstrom, I., et al., *A Swedish register-based, long-term inception cohort study of patients with rheumatoid arthritis - results of clinical relevance.* Open Access Rheumatol, 2019. **11**: p. 207-217.
23. de Brito Rocha, S., D.C. Baldo, and L.E.C. Andrade, *Clinical and pathophysiologic relevance of autoantibodies in rheumatoid arthritis.* Adv Rheumatol, 2019. **59**(1): p. 2.
24. Fert-Bober, J., E. Darrah, and F. Andrade, *Insights into the study and origin of the citrullinome in rheumatoid arthritis.* Immunol Rev, 2020. **294**(1): p. 133-147.
25. Prevoo, M.L., et al., *Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis.* Arthritis Rheum, 1995. **38**(1): p. 44-8.
26. Andersson, M.L.E., et al., *Patients with Early Rheumatoid Arthritis in the 2000s Have Equal Disability and Pain Despite Less Disease Activity Compared with the 1990s: Data from the BARFOT Study over 8 Years.* J Rheumatol, 2017. **44**(6): p. 723-731.
27. Glaser, C., et al., *Whipple's disease mimicking rheumatoid arthritis can cause misdiagnosis and treatment failure.* Orphanet J Rare Dis, 2017. **12**(1): p. 99.
28. Benjamanukul, S., et al., *Rheumatic manifestations of Chikungunya virus infection: Prevalence, patterns, and enthesitis.* PLoS One, 2021. **16**(4): p. e0249867.
29. Moore, T.L., *Parvovirus-associated arthritis.* Curr Opin Rheumatol, 2000. **12**(4): p. 289-94.
30. Jacobs, J.W., *Optimal use of non-biologic therapy in the treatment of rheumatoid arthritis.* Rheumatology (Oxford), 2012. **51 Suppl 4**: p. iv3-8.
31. Hetland, M.L., et al., *Active conventional treatment and three different biological treatments in early rheumatoid arthritis: phase IV investigator initiated, randomised, observer blinded clinical trial.* BMJ, 2020. **371**: p. m4328.
32. Boeters, D.M., et al., *The 2010 ACR/EULAR criteria are not sufficiently accurate in the early identification of autoantibody-negative rheumatoid arthritis: Results from the Leiden-EAC and ESPOIR cohorts.* Semin Arthritis Rheum, 2017. **47**(2): p. 170-174.
33. Vonkeman, H.E. and M.A. van de Laar, *The new European League Against Rheumatism/American College of Rheumatology diagnostic criteria for rheumatoid arthritis: how are they performing?* Curr Opin Rheumatol, 2013. **25**(3): p. 354-9.

Populärvetenskaplig sammanfattning

Examensarbete

Läkarprogrammet, Avd för reumatologi och inflammationsforskning 2021

Författare: Sophie Hallberg

Supervisor: PhD Monica Leu Agelii

Rätt eller fel diagnos – vem har egentligen ledgångsreumatism?

Ledgångsreumatism är en ganska vanlig åkomma som drabbar ungefär 1% av befolkningen. Det innebär att Sverige finns det ca 100 000 personer med sjukdomen, och de flesta som drabbas är medelålderskvinnor. Som namnet antyder drabbas framförallt kroppens leder och typiskt är att det börjar i fingrarnas småleder. Symptomen är smärta, svullnad, svårigheter att röra sig, stelhet och en stor trötthet. Ledgångsreumatism beror på att kroppens immunförsvar, som normalt skall skydda oss mot virus och bakterier, har fått för sig att lederna är något främmande och attackerar dem. Ledgångsreumatism är en obotbar men fullt behandlingsbar sjukdom och behandlingen går ut på att dämpa aktiviteten i immunförsvaret utan att för den skulle göra patienten känslig för infektioner, en svår balansgång. Behandlar man inte så kommer lederna att brytas ned vilket leder till handikapp och dessutom kan även inre organ som t ex njurar angripas. Idag är medicinerna bra, men de är starka och kan ha allvarliga biverkningar. Dessutom är de dyra. Det gäller därför att försöka ge rätt patient rätt sorts medicin och för att göra det måste vi kunna ställa rätt diagnos tidigt i sjukdomsförloppet.

Att ställa diagnosen ledgångsreumatism är komplicerat. Det finns inte något enskilt symptom eller något blodprov som helt säkert säger att detta är ledgångsreumatism. Istället använder man sig av olika kriterier. Idag är det viktigt att tidigt hitta patienter med diagnosen eftersom vi vet att tidig behandling har bättre effekt. Men tidig diagnos kan ibland bli fel diagnos och då kan patienten få fel, för lite eller för mycket behandling. I denna studien har vi undersökt om de kriterier som vi använt för att diagnosticera ledgångsreumatism i en större studie är tillräckligt bra för att ställa diagnosen eller om diagnosen kommer ändra sig över tid.

Vi har undersökt en grupp av ca 2500 patienter som fått diagnosen ledgångsreumatism och sedan följts under 5 års tid. Alla dessa patienter uppfyllde 4 av de 7 kriterier (stelhet på morgonen, inflammation i 3 eller fler leder, inflammation i händerna, inflammation både i höger och vänster kroppshalva, reumatisk faktor i blodet, bindvävsknutor eller förstörda leder som man ser på röntgen) som man 1987 hade bestämt vara typiska för sjukdomen, dvs de hade ledgångsreumatism.

När vi efter 5 år tittade efter om patienterna fortfarande hade diagnosen ledgångsreumatism såg vi att 44 stycken, det innbär knappt 2% av alla patienter, hade bytt diagnos. Så det var ganska bra träffsäkerhet. Det som framförallt skiljde de som hade kvar sin diagnos var att de uppfyllde fler än de nödvändiga 4 kriterierna för diagnosen och av dessa kriterier så såg vi de oftast hade markören reumatisk faktor i blodet och de hade röntgenförändringar i sina leder.

Så om en patient som kommer med svullna leder inte har någon reumatisk faktor, inga röntgenförändringar och endast uppfyller 4 av kriteriena skall man fundera på om detta verkligen är ledgångsreumatism. Det blir också viktigt att vid varje återbesök fundera över om diagnosen när rätt särskilt om inte patienten svarar som förväntat på behandlingen.

Acknowledgements

Inger Gjertsson: I am forever extremely grateful for your encouragement!

PhD Monica Leu Agelii: I am equally grateful for your support as my supervisor.

Alexander Strid Holmertz: My dear friend! You know what you have meant for this degree project! I am so thankful!

Mats Dehlin: Thankyou for taking the time with my examination!

My uncle Urban Hallberg and my mum Helena Hjortander: Thankyou for all your encouragement and patience during this project.