Different Viewpoints on Tapering DMARDs in Rheumatoid Arthritis

Elise van Mulligen

Different Viewpoints on Tapering DMARDs in Rheumatoid Arthritis

Verschillende perspectieven op het afbouwen van DMARDs bij reumatoïde artritis

Elise van Mulligen

Cover design: Publiss

Lay-out: Publiss | www.publiss.nl

Print: Ridderprint | www.ridderprint.nl.

© Copyright 2021: Elise van Mulligen

All rights reserved. No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without prior permission of the author.

Financial support for the publication of this thesis was kindly provided by the Erasmus University Medical Center, Pfizer B.V., Eli Lilly Nederland B.V., and Galapagos B.V..

Different Viewpoints on Tapering DMARDs in Rheumatoid Arthritis

Verschillende perspectieven op het afbouwen van DMARDs bij reumatoïde artritis

Proefschrift

ter verkrijging van de graad van doctor aan de

Erasmus Universiteit Rotterdam

op gezag van de

rector magnificus

Prof. dr. F.A. van der Duijn Schouten

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

dinsdag 2 maart 2021 om 15:30 uur

door

Elise van Mulligen

geboren te Utrecht.

Erasmus University Rotterdam

Promotiecommissie:

Promotoren:	prof. dr. J.M.W. Hazes prof. dr. A.H.M. van der Helm - van Mil prof. dr. A.E.A.M. Weel
Overige leden:	prof. dr. A.H.J. Mathijssen prof. dr. C.J. van der Woude dr. A.A. den Broeder
Copromotor:	dr. P.H.P. de Jong

Table of contents

CHAPTER 1

INTRODUCTION

Management of Rheumatoid Arthritis

Treatment outcomes in rheumatoid arthritis (RA) have improved enormously in the last decades due to early initiation of therapy, a treat-to-target approach, and use of biological disease-modifying anti-rheumatic drugs (DMARDs).[1] A treat-to-target strategy, aiming at a pre-set outcome measure, is well-embedded in daily practice for management of RA. The ultimate aim of every treat-to-target strategy is reaching sustained remission, defined as the total absence of both articular and extra-articular manifestations of RA, with low disease activity as the best alternative. This is obtained with treatment intensifications until the target is reached. This treatment strategy has the highest chance of inducing remission and preventing joint damage.[2, 3]

As a result, remission in RA occurs more often.[4] Approximately 75%-80% of patients are able to reach low disease activity or even sustained remission.[4] This has raised the question whether we need to continue, taper or discontinue treatment. Reasons for tapering are reduction in costs, since treatment with biologicals is very expensive, prevention of possible long term side effects, and patient preference. However, by tapering medication the risk of disease flares increases, which can have a great impact on patients lives, and on society due to productivity loss, i.e. sick leave or unemployment. Therefore, before tapering treatment is considered, it is important to evaluate both the clinical-, patient-, and societal viewpoint to make informed decisions. Moreover, the optimal tapering strategy leading to the least amount of disease flares has not been developed yet.

Clinical perspective

Management of RA must be both effective and affordable, which can be accomplished by efficient use of DMARDs. Therefore, when RA patients have a well-controlled disease, and reach a symptom-free state, continuation of all DMARDs is no longer required. Especially since the benefits of treatment should always outweigh possible side-effects, which may not be the case when RA patients have a well-controlled disease. Long term use of biologicals, for example, is not without risks and includes adverse events related to the immunosuppressive mode of action, such as infections and malignancies.[5, 6]

Although rheumatologists carefully consider initiation of biologicals, uniform tapering decisions are lacking. As a result, tapering of biologicals is not common practice in RA

patients with a well-controlled disease.[7] A possible reason behind this trend is the increased risk of disease flares during tapering. Previous studies, in which biologicals were most frequently discontinued instead of gradually tapered, have shown that it is possible to withdraw treatment, but this is accompanied with a higher chance at disease flares. Flare rates within these studies varied from 38-76.6%.[8-11]

To optimize treatment for symptom-free patients, a more tailored treatment approach is needed. Building further on a treat-to-target strategy with disease activity as outcome parameter, tapering treatment should be performed as soon as patients reach well-controlled disease. Current EULAR treatment guidelines recommend that rheumatologists should consider tapering of biologicals when a RA patient is in sustained remission, especially when combined with a csDMARD. If a patient is in remission with only a csDMARD, the csDMARD may be tapered.[1] However, neither remission criteria are defined nor a time interval is given for the term sustained. Furthermore, the best method to taper treatment still needs to be unraveled.

Patient perspective

Tapering treatment implies less medication use, which is often preferred by patients. It also decreases the amount of burden due to subcutaneous or intravenous administration of biologicals.

However, the main concern for patients is that tapering of treatment will lead to a disease flare, which could lead to more pain and disability.[8, 12-14] It has also been shown that only 41 – 67% of the patients that experience a flare will regain remission within 6 months after treatment intensification.[8, 15] Thus, for some patients tapering seems feasible, while for others it could lead to a flare followed by a reduced or no response to previous effective therapy.

Nowadays, a paradigm shift in the delivery of health care is emerging and is shifting towards patient centered health care. Patient centered healthcare focuses on the individual patient preferences and needs, which can be objectified with patient reported outcomes (PROs).[16, 17] In order to optimize the delivery of care during tapering we need to know the consequences of a flare, which could be measured with these PROs. A previous study already showed that discontinuation of treatment has a short significant impact on PROs.[18] However, data on the effect of a disease flare on patients' lives are sparse, while this is one of the main concerns of patients.

Societal perspective

One of the main reasons for tapering treatment is saving costs, especially since health care costs are rising due to the use of biological treatment. Annually, more than 250 million euros are spent on the use of biological therapy for rheumatic diseases in the Netherlands.[19] To keep health care affordable, tapering treatment in RA patients with a well-controlled disease could lead to considerable cost-savings. However, the costs that are saved by tapering biologicals may be counterbalanced by increased costs due to productivity loss as a result of the higher flare rates during tapering of treatment. Currently, it is unclear whether health care costs outweigh the possible increase in societal costs, especially because recent cost-effectiveness analyses only take into account sick leave (absenteeism), while working while being sick (presenteeism), is neglected.

Ultimate treatment outcome: DMARD-free remission?

If we are able to safely taper and discontinue medication in RA patients who have a well-controlled disease, patients will reach a state of DMARD-free remission (DFR; the absence of synovitis after cessation of DMARD therapy). The ability to achieve and sustain DFR is often considered unlikely.[20] Nonetheless, there is increasing interest in achieving DFR, because this is currently the best proxy of a cure for RA.[21] Previous research showed that 10-20% of RA patients are able to achieve sustained DFR.[21, 22] However, definitions for DFR are heterogeneous. Although, DFR has been mentioned as an outcome for early RA patients, there is no to little evidence that DFR is also achievable in an established RA population.

TARA trial

The TApering strategies in Rheumatoid Arthritis (TARA) trial was set-up to investigate the best tapering order for RA patients with a well-controlled disease who were using both a TNF-inhibitor and a csDMARD. The TARA, a multicenter, single-blinded (research nurses), randomised trial was carried out in twelve rheumatology centers in the south-western part of the Netherlands. Inclusion started in September 2011 and ended in July 2016. Adult RA patients with a well-controlled disease, defined as a disease activity score (DAS) ≤ 2.4 and a swollen joint count (SJC) ≤ 1 for more than three months, using a combination of a csDMARD and TNF-inhibitor, were included. Patients were randomised into gradual tapering either the csDMARD in the first year

followed by the TNF-inhibitor in the second year, or vice versa. csDMARD tapering was realized by cutting the dosage into half, a quarter and thereafter it was stopped. The TNF-inhibitor was tapered by doubling the dose interval, followed by cutting the dosage into half, and thereafter it was stopped. The total tapering schedule took six months, with dose adjustments every three months as long as there was still a well-controlled disease. If a disease flare occurred, defined as DAS > 2.4 and/or SJC > 1, tapering was stopped and the last effective treatment was restarted and if necessary, medication was intensified further according to a treat-to-target approach, until low disease activity was reached again. After a flare, no further attempts were taken to taper medication during the remainder of the study.

Patients were examined at baseline and every three months thereafter. At each time point, the DAS, medication usage, and self-reported questionnaires were collected, except for hand and foot radiographs, which were obtained at baseline and after one and two years of follow-up. Throughout the whole study follow-up adverse event were recorded.

Objectives and outline of this thesis

More and more RA patients will reach remission with current improved treatment strategies. As a result, tapering of treatment will become more common, but the optimal tapering strategy leading to the least amount of flares, has not been developed yet. Furthermore, the impact of tapering and flares on a patient and a societal level is currently not well known.

To objectify the possible impact of tapering in daily practice, we described current biological use in the Netherlands in chapter 2. For this we used a real-world observational cohort in which we investigated factors influencing biological survival, thereby taking into account various reasons for discontinuation.

In chapter 3 we present the first year results of the aforementioned TARA trial. We compared two gradual tapering strategies, namely tapering the TNF-inhibitor first followed by the csDMARD and vice versa. We assessed which therapy should be tapered first based on the tapering strategy leading to the least amount of flares. In chapter 4 additional information is given on the treatment strategies that patients were using before inclusion, which is also an indicator for the treatment strategies that are currently used in daily practice.

When treatment is completely tapered and discontinued, a patient will reach the state of DMARD-free remission. In chapter 5 we present a systematic literature review investigating the feasibility of DMARD-free remission as a novel and sustainable outcome for RA.

Following this, in chapter 6 we investigated whether DMARD-free remission is also an achievable treatment outcome in an established RA population. In chapter 7 we compared data from the TARA trial to real-world tapering data.

In chapter 8 we determined the impact of a disease flare on patient's lives. We took into account patient reported outcomes (PROs) and investigated whether these PROs changed if a flare occurred, and if so, the duration of this effect was determined.

An important reason for tapering treatment is to save costs. In chapter 9 we investigated the societal impact of tapering, in which we evaluated the cost-effectiveness of the two tapering strategies in the TARA trial. Health care costs as well as societal costs, i.e. costs due to loss of productivity, are taken into account to decide which tapering strategy is most cost-effective.

Last, in chapter 10 a general discussion is provided of the main findings of this thesis and their implications for current clinical practice. Finally, suggestions for future research are discussed.

References

1. Smolen JS, Landewe RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, McInnes IB, Sepriano A, van Vollenhoven RF, de Wit M *et al*: EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. *Ann Rheum Dis* 2020.
2. Klarenbeek NB, Allaart CF, Kerstens PJ, Huizinga TW, Dijkmans BA: The BeSt story: on strategy trials in rheumatoid arthritis. *Curr Opin Rheumatol* 2009, 21(3):291-298.
3. de Jong PH, Hazes JM, Han HK, Huisman M, van Zeben D, van der Lubbe PA, Gerards AH, van Schaeybroeck B, de Sonnaville PB, van Krugten MV *et al*: Randomised comparison of initial triple DMARD therapy with methotrexate monotherapy in combination with low-dose glucocorticoid bridging therapy; 1-year data of the tREACH trial. *Ann Rheum Dis* 2014, 73(7):1331-1339.
4. Aletaha D, Smolen JS: Diagnosis and Management of Rheumatoid Arthritis: A Review. *Jama* 2018, 320(13):1360-1372.
5. Ramiro S, Sepriano A, Chatzidionysiou K, Nam JL, Smolen JS, van der Heijde D, Dougados M, van Vollenhoven R, Bijlsma JW, Burmester GR *et al*: Safety of synthetic and biological DMARDs: a systematic literature review informing the 2016 update of the EULAR recommendations for management of rheumatoid arthritis. *Ann Rheum Dis* 2017, 76(6):1101-1136.
6. Pereira R, Lago P, Faria R, Torres T: Safety of Anti-TNF Therapies in Immune-Mediated Inflammatory Diseases: Focus on Infections and Malignancy. *Drug Dev Res* 2015, 76(8):419-427.
7. Kuijper TM, Folmer R, Stolk EA, Hazes JMW, Luime JJ: Doctors' preferences in de-escalating DMARDs in rheumatoid arthritis: a discrete choice experiment. *Arthritis Res Ther* 2017, 19(1):78.
8. Ghiti Moghadam M, Vonkeman HE, Ten Klooster PM, Tekstra J, van Schaardenburg D, Starmans-Kool M, Brouwer E, Bos R, Lems WF, Colin EM *et al*: Stopping Tumor Necrosis Factor Inhibitor Treatment in Patients With Established Rheumatoid Arthritis in Remission or With Stable Low Disease Activity: A Pragmatic Multicenter, Open-Label Randomized Controlled Trial. *Arthritis Rheumatol* 2016, 68(8):1810-1817.
9. Fautrel B, den Broeder AA: De-intensifying treatment in established rheumatoid arthritis (RA): Why, how, when and in whom can DMARDs be tapered? *Best Pract Res Clin Rheumatol* 2015, 29(4):550-565.
10. van Herwaarden N, den Broeder AA, Jacobs W, van der Maas A, Bijlsma JW, van Vollenhoven RF, van den Bemt BJ: Down-titration and discontinuation strategies of tumor necrosis factor-blocking agents for rheumatoid arthritis in patients with low disease activity. *Cochrane Database Syst Rev* 2014(9):CD010455.
11. Haschka J, Englbrecht M, Hueber AJ, Manger B, Kleyer A, Reiser M, Finzel S, Tony H-P, Kleinert S, Feuchtenberger 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* 2016, 75(1):45-51.
12. Kuijper TM, Luime JJ, de Jong PH, Gerards AH, van Zeben D, Tchetverikov I, de Sonnaville PB, van Krugten MV, Grillet BA, Hazes JM *et al*: Tapering conventional synthetic DMARDs in patients with early arthritis in sustained remission: 2-year follow-up of the tREACH trial. *Ann Rheum Dis* 2016, 75(12):2119-2123.

13. Schett G, Emery P, Tanaka Y, Burmester G, Pisetsky DS, Naredo E, Fautrel B, van Vollenhoven R: Tapering biologic and conventional DMARD therapy in rheumatoid arthritis: current evidence and future directions. *Ann Rheum Dis* 2016, 75(8):1428-1437.
14. Verhoef LM, Selten EMH, Vriezekolk JE, de Jong AJL, van den Hoogen FHJ, den Broeder AA, Hulscher ME: The patient perspective on biologic DMARD dose reduction in rheumatoid arthritis: a mixed methods study. *Rheumatology (Oxford)* 2018.
15. Fautrel B, Pham T, Alfaiate T, Gandjbakhch F, Foltz V, Morel J, Dernis E, Gaudin P, Brocq O, Solau-Gervais E *et al*: Step-down strategy of spacing TNF-blocker injections for established rheumatoid arthritis in remission: results of the multicentre non-inferiority randomised open-label controlled trial (STRASS: Spacing of TNF-blocker injections in Rheumatoid ArthritiS Study). *Ann Rheum Dis* 2016, 75(1):59-67.
16. Heller JE, Shadick NA: Outcomes in rheumatoid arthritis: incorporating the patient perspective. *Curr Opin Rheumatol* 2007, 19(2):101-105.
17. Oude Voshaar MAH, Das Gupta Z, Bijlsma JWJ, Boonen A, Chau J, Courvoisier DS, Curtis JR, Ellis B, Ernestam S, Gossec L *et al*: The International Consortium for Health Outcome Measurement (ICHOM) Set of Outcomes that Matter to People Living with Inflammatory Arthritis Consensus from an international Working Group. *Arthritis Care Res (Hoboken)* 2018.
18. Ghiti Moghadam M, Ten Klooster PM, Vonkeman HE, Kneepkens EL, Klaasen R, Stolk JN, Tchetverikov I, Vreugdenhil SA, van Woerkom JM, Goekoop-Ruiterman YPM *et al*: Impact of Stopping Tumor Necrosis Factor Inhibitors on Rheumatoid Arthritis Patients' Burden of Disease. *Arthritis Care Res (Hoboken)* 2018, 70(4):516-524.
19. Biologicals, 2014-2018 [https://www.gipdatabank.nl/]
20. O'Mahony R, Richards A, Deighton C, Scott D: Withdrawal of disease-modifying antirheumatic drugs in patients with rheumatoid arthritis: a systematic review and meta-analysis. *Ann Rheum Dis* 2010, 69(10):1823-1826.
21. Ajeganova S, van Steenbergen HW, van Nies JA, Burgers LE, Huizinga TW, van der Helm-van Mil AH: Disease-modifying antirheumatic drug-free sustained remission in rheumatoid arthritis: an increasingly achievable outcome with subsidence of disease symptoms. *Ann Rheum Dis* 2016, 75(5):867-873.
22. Akdemir G, Heimans L, Bergstra SA, Goekoop RJ, van Oosterhout M, van Groenendael J, Peeters AJ, Steup-Beekman GM, Lard LR, de Sonnaville PBJ *et al*: Clinical and radiological outcomes of 5-year drug-free remission-steered treatment in patients with early arthritis: IMPROVED study. *Ann Rheum Dis* 2018, 77(1):111-118.

CHAPTER 2

FACTORS THAT INFLUENCE BIOLOGICAL SURVIVAL IN RHEUMATOID ARTHRITIS: RESULTS OF A REAL-WORLD COHORT FROM THE NETHERLANDS

E van Mulligen, S Ahmed, AEAM Weel, JMW Hazes, AHM van der Helm- van Mil, PHP de Jong

Submitted

ABSTRACT

Objectives We aim to explore real-world biological survival stratified for discontinuation reason and determine its influenceability in rheumatoid arthritis(RA) patients.

Methods Data from the local pharmacy database and patient records of a university hospital in the Netherlands were used. RA patients who started a biological between 2000-2020 were included. Data on age, anti-citrullinated protein antibody (ACPA) and rheumatoid factor (RF)-status, presence of erosions, gender, body mass index, time to first biological, biological survival time, use of csDMARDs and discontinuation reasons were collected.

Results Of the included 318 patients, 12% started their first biological within 6 months after diagnosis. Median time to first biological was 3.6 years (95%CI,1.0-7.2). Median survival of the first- and second-line biological was respectively 1.7 years (95%CI,1.3-2.2) and 0.8 years (95%CI, 0.5-1.0)(p=0.0001). Discontinuation reasons for the first-line biological were ineffectiveness(47%), adverse events(17%), remission(16%), pregnancy(30%), or patient preference(10%). Multivariable Cox regression analyses for discontinuation due to inefficacy or adverse events showed that concomitant use of csDMARDs (HR=1.32,p<0.001) positively, while RF-positivity negatively (HR=0.82,p=0.03) influenced biological survival. ACPA-positivity was associated with longer biological survival due to inability to discontinue due to remission (HR=1.43,p=0.023). Second-line TNF-inhibitor survival was similar between patients with a primary and secondary non-response on the first-line TNF-inhibitor (HR=1.28,p=0.34).

Conclusion Biological survival diminishes with the number of biologicals used. Biological survival is prolonged if patients use csDMARDs. RF and ACPA were negatively associated with respectively biological survival and discontinuation due to remission. Therefore, tailoring treatment based upon autoantibody status might be the first step towards personalized medicine in RA.

INTRODUCTION

Management of RA has improved in the last decades due to early diagnosis, a treat-to-target approach, and the introduction of biological disease modifying anti-rheumatic drugs(bDMARDs).[1] Tumor necrosis factor inhibitors (TNF-inhibitors) were the first bDMARDs to be developed for rheumatic diseases and are currently most frequently prescribed after an inadequate response to conventional synthetic (cs)DMARDs. It has been suggested that prolonged biological survival is a surrogate for treatment effectiveness.[2] However, an increasing amount of patients reach remission nowadays, and will taper and discontinue treatment.[1] Therefore, solely taking into account overall biological survival will dilute outcomes, and to properly analyze biological survival, results should be stratified according to discontinuation reasons.

Previous studies, based on biological registries throughout Europe, have shown that 50% of patients discontinue their TNF-inhibitor within 3-5 years.[3] Main reasons for discontinuation were inefficacy and adverse events.[3, 4] Within trials and biological registries longer survival times were seen for first-line biologicals and when bDMARDS were combined with csDMARDs.[5-7] However, factors influencing biological survival based on separate reasons for discontinuation have not been previously explored.

Therefore, the aim of this Dutch real-world rheumatoid arthritis cohort is to explore first and second-line biological survival and to determine its influenceability when stratified for discontinuation reasons.

PATIENTS AND METHODS

Study design

Data from the local pharmacy database and patient records of the Erasmus MC, an academic hospital in the Netherlands, were used. We included data from rheumatoid arthritis(RA) patients starting a biological between 2000-2020. We excluded patients for whom non-adherence was reported, and if start and stop dates for bDMARDs were not available. Standard treatment of RA in the Netherlands is based upon a treat-to-target approach aiming for low disease activity. If patients have an inadequate response to >1 csDMARD, a bDMARD can be prescribed. In case of an inadequate

response, rheumatologists can prescribe another TNF-inhibitor(cycling) or a bDMARD with another mode of action(switching).[8]

Data collection

Biological survival was the main outcome and was defined as skipping ≥2 gifts and/or ≥2 months without biological treatment. Reasons for discontinuations were evaluated and classified into: inefficacy; adverse events (AEs), which we divided into primary (<6 months) and secondary (≥6 months) non-response; remission; pregnancy; patient preference; and other reasons.

Analyses

We compared first- and second-line biological survival with Kaplan-Meier curves and with Wilcoxon-Breslow-Gehan tests at 3 years. Thereafter, first-line biological survival with and without concomitant use of csDMARD(s) was compared. Subsequently, we investigated whether primary and secondary inefficacy to a first-line TNF-inhibitor leads to differences in second-line TNF-inhibitor survival. Patients stopping their bDMARD due to remission or pregnancy were censored.

Cox proportional hazard models were used to estimate hazard ratios (HRs) of candidate baseline predictors (age, gender, ACPA, RF, erosions, BMI, disease duration, or co-medication) for bDMARD survival stratified for reasons for discontinuation, namely (1) inefficacy or adverse events and (2) remission. First univariable Cox regression analyses were performed, and candidate predictors with a $p<0.20$ were entered into a multivariable model, after which backward selection was applied until significance was reached. To prevent overfitting, an entry model was created and backward selection was applied. Schoenfeld residuals were assessed to check the proportional hazard assumption.

All data was analyzed using STATA15. P-values ≤0.05 were considered statistically significant.

RESULTS

Patients

Data were derived from 318 RA patients(table 1). Time until first bDMARD prescription remained constant between 2000 and 2020. In our cohort 50% of patients started their first biological after 2013, thus in most recent years more bDMARDs were prescribed. A total of 39 (12%) patients started their first bDMARD within 6 months after diagnosis.

Table 1 Characteristics of rheumatoid arthritis population using a biological in a university hospital

	RA patients, n=318
Demographic	
- Age at diagnosis, mean (sd)	40.9 (16)
- Gender, female, n (%)	264 (83)
- BMI, mean (sd)	26.9 (6.3)
Disease characteristics	
- ACPA positive, n (%)	224 (70)
- RF positive, n (%)	226 (71)
- Erosive disease, n (%)	141 (44)
Medication	
- Time to first biological, years, median (IQR)	3.6 (1-7)
- First-line biologicals	
• Etanercept, n (%)	142 (45)
• Adalimumab, n (%)	90 (28)
• Certolizumab Pegol, n (%)	59 (19)
• Infliximab, n (%)	15 (5)
• Golimumab, n (%)	5 (2)
• Anakinra, n (%)	3 (1)
- csDMARDs used with first-line biological	
• MTX, n (%)	66 (21)
• MTX + SASP and/or HCQ , n (%)	147 (46)
• Other csDMARDs (SASP, HCQ, LEF), n (%)	53 (17)
• No combination therapy, n (%)	52 (16)

ACPA: anti-citrillunated protein antibody, BMI: body mass index, csDMARD: conventional synthetic disease modifying anti-rheumatic drug, HCQ: hydroxychloroquine, IQR: inter quartile range, LEF: leflunomide, MTX: methotrexate, RF: rheumatoid factor, SASP: sulfasalazine, sd: standard deviation

First- and second-line biological survival

Median (95%CI) survival time of the first-line biological was 1.7 years (1.3-2.2), and for the second-line bDMARD 0.8 years (0.5-1). Most prescribed first-line bDMARDs were Etanercept (45%), Adalimumab (28%), and Certolizumab Pegol (19%)(table 1). Since only 9% of patients were using non-TNF-inhibitors as second-line bDMARD, a direct comparison between a cycling or switching strategy could not be performed.

bDMARD survival was significantly longer for the first-line bDMARD compared to the second (p=0.0001)(figure 1A). Discontinuation reasons for the first-line bDMARD were inefficacy (47%), adverse events (17%), remission (16%), pregnancy (30%), or patient preference (10%). Discontinuation reasons for the second-line bDMARD were similar (supplemental table S1).

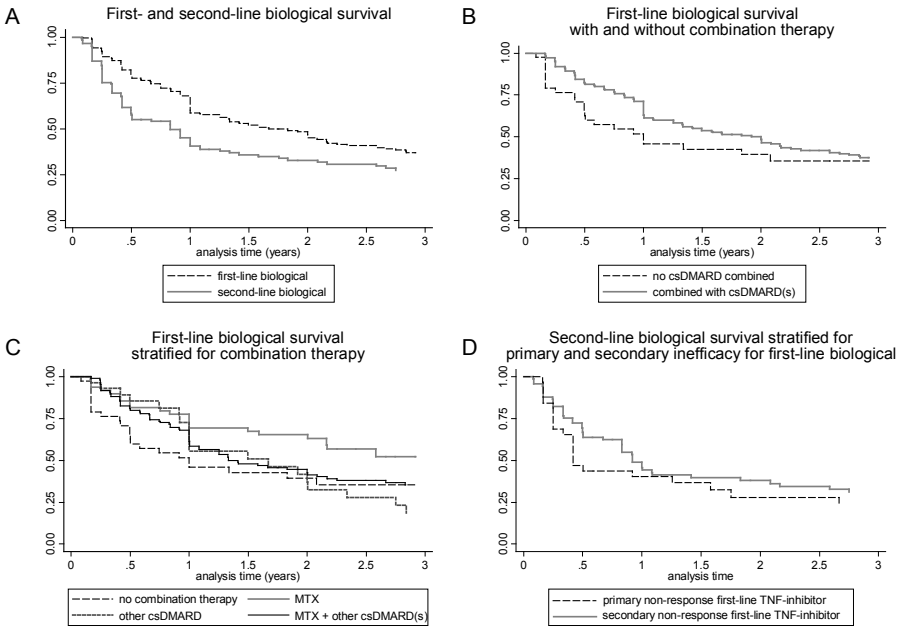


Figure 1 Kaplan-Meier curves for biological survival. (A) Kaplan-Meier for first- versus second-line biological survival, (B) Kaplan-Meier curve of patients with or without combination therapy, (C) Kaplan-Meier curve of patients without combination therapy, and for patients with combination therapy stratified for methotrexate, methotrexate combined with one or more other csDMARDs (sulfasalazine, hydroxychloroquine, and/or leflunomide), or one or more other csDMARDs and (D) Kaplan-Meier of second-line TNF-inhibitor survival, stratified for primary and secondary inefficacy for the first-line TNF-inhibitor.

csDMARD: conventional synthetic disease modifying anti-rheumatic drugs, MTX: methotrexate.

First-line biological survival with or without concomitant use of csDMARDs

A total of 48(25.3%) and 6(15.4%) patients respectively with and without concomitant use of csDMARD(s) were still using their first-line biological after 3 years of follow-up. Median (95%CI) survival time of the first-line bDMARD with csDMARD(s) was 2.0 (1.3-2.3) years, and without csDMARDs 1.0 (0.5-5.3) year (figure 1B,p=0.031). First-line bDMARD survival was longest for treatment regimens with methotrexate (MTX) followed by other csDMARDs, and no csDMARD use (figure 1C). However, no significant differences were found between MTX and the other csDMARDs as concomitant therapy (p=0.14)(figure 1C).

Primary and secondary failure

Median (95%CI) survival time for the second-line TNF-inhibitor was 0.42(0.25-1.58) years for patients with a primary non-response for the first TNF-inhibitor and 0.92(0.83-1.83) years for patients with a secondary non-response for the first TNF-inhibitor. Although overall survival time on the second-line biological did not differ significantly between patients with a primary and secondary non-response (HR 1.28,p=0.34), a trend could be observed (figure 1D).

Predictors for biological survival

Univariate cox regression for discontinuation due to inefficacy and adverse events showed that RF (HR=0.80,p=0.014), and presence of erosions(HR=0.65,p<0.001) were negatively associated with first-line bDMARD survival. Concomitant use of csDMARD(s) (HR=1.35,p<0.001) on the other hand was positively associated with first-line bDMARD survival. Aforementioned factors as well as time to first-line bDMARD, age, gender and ACPA were included in our multivariable model with backward selection. In the final model only RF(HR=0.82,p=0.03), and concomitant use of csDMARDs(HR=1.32,p=0.001) were significantly associated with first-line bDMARD survival (table 2). When we used an entry model and applied backward selection, aforementioned predictors were again in the final model, but also the presence of erosions was included.

The same procedure was followed for investigating which factors were associated with a higher chance of discontinuing bDMARDs due to remission. Only a positive ACPA-status was associated with longer biological survival due to inability to taper medication(HR=1.43,p=0.023)(table 2).

Table 2 Predictors for biological survival

	Univariable		Multivariable[1]	
	HR (95% CI)	**p**	**HR (95% CI)**	**p**
Biological survival taking into account discontinuation due to inefficacy or AEs[2]				
Age at diagnosis	1.00 (1.00-1.01)	0.514		
Gender (female)	1.00 (0.82-1.23)	0.985		
BMI	0.99 (0.98-1.01)	0.296		
Rheumatoid factor	**0.80 (0.67-0.96)**	**0.014**	**0.82 (0.69-0.98)**	**0.03**
ACPA	0.90 (0.75-1.07)	0.223		
Erosions	**0.65 (0.55-0.76)**	**<0.001**		
Time to first-line biological	**0.98 (0.95-1.01)**	**0.163**		
Combination therapy	**1.35 (1.14-1.59)**	**<0.001**	**1.32 (1.13-1.57)**	**0.001**
Prolonged biological survival due to inability to taper[3]				
Age at diagnosis	1.00 (0.99-1.01)	0.717		
Gender (female)	1.08 (0.75-1.56)	0.676		
BMI	**0.98 (0.95-1.01)**	**0.175**		
Rheumatoid factor	**1.26 (0.94-1.96)**	**0.121**		
ACPA	**1.43 (1.05-1.93)**	**0.023**	**1.43 (1.05-1.93)**	**0.023**
Erosions	0.70 (0.53-0.92)	0.481		
Time to first-line biological	**1.04 (0.99-1.09)**	**0.119**		
Combination therapy	0.93 (0.69-1.26)	0.643		

[1]Backward selection, variables with $p<0.20$ in univariable analyses were entered. [2]HR>1 indicates prolonged biological survival, HR<1 indicates reduced biological survival due to inefficacy or AEs. [3]HR>1 indicates prolonged biological survival due to inability to taper, HR<1 indicates reduced biological survival due to tapering of bDMARD due to remission.

ACPA: anti-citrullinated protein antibody, AE: adverse events, BMI: body mass index, CI: confidence interval, HR: hazard ratio.

DISCUSSION

Optimal management of RA is based on reaching the lowest possible disease activity with a treat-to-target approach.[1] Despite the improved management approach and increasing treatment options, only 60-70% of RA patients will reach a long-term clinical response.[4] Within our study we found a significant difference in survival time between the first- and second-line bDMARD, implicating the importance to prolong first-line bDMARD survival. Several factors can influence bDMARD survival of which some can be influenced.

Main reasons for discontinuation in our and in other studies were inefficacy and adverse events.[3] Primary inefficacy indicates no effect at all, and is thought to be due to a mismatch between the bDMARD and the specific RA-subtype, causing the biologic agent not to be effective.[9] Secondary inefficacy indicates that the clinical response is first obtained, but not maintained, and is thought to be caused by formation of auto-antibodies against the biologic.[4] Although we did not find a significant difference in second-line TNF-inhibitor survival between RA patients with a primary or secondary non-response to the first TNF-inhibitor, a trend could be observed. This was probably due to a low number of patients in the group with a primary non-response for the first-line bDMARD (n=42). However, these data indicate that rheumatologists should consider to switch to another mode of action in case of primary inefficacy instead of cycling to another TNF-inhibitor, but validation is needed.[10, 11]

Compared to previous findings, bDMARD survival seems to be short. This can be explained by our real-life cohort in a tertiary care university hospital. We also noticed that in our cohort a high number of patients were using Certolizumab Pegol, and discontinued their bDMARD due to pregnancy. This is related to the fact that the ErasmusMC has an ongoing cohort for patients with a wish to conceive.[12]

Outcomes of our study on the other hand are in accordance with previous findings. Benefits of combining a bDMARD with a csDMARD have been previously described. [6, 13, 14] Reasons for this synergistic effect are not fully understood. One of the reasons could be that csDMARDs can prevent development of neutralizing anti-drug-antibodies. It is also thought that csDMARDs affect clearance of the bDMARD by modulating either the expression of Fc receptors on monocytes or the interaction of the Fc receptor and the bDMARD.[4]

RF positivity was also found to be predictive for shorter bDMARD survival, which is again in accordance with previous literature.[15] Furthermore, ACPA positivity has shown to reduce the chance of discontinuing due to remission. This reconfirms that the presence of autoantibodies are a marker for more severe disease.

In conclusion, bDMARD survival diminishes with the number of bDMARDs used. Combining a bDMARD with a csDMARD increases bDMARD survival, which supports current EULAR recommendations to combine a bDMARD with a csDMARD. RF and ACPA were negatively associated with respectively bDMARD survival and discontinuation due to remission. Therefore, the possible first step to personalized medicine in RA might be tailoring of treatment based upon autoantibody status.

Table S1 Discontinuation reasons for first- and second-line biological

	First-line biological (n=318)	**Second-line biological (n=192)**
Total number of patients discontinuing biological treatment	226 (71)	127 (66)
Ineffective	106 (47)	63 (51)
Adverse event	38 (17)	28 (23)
Remission	35 (16)	13 (10)
Pregnancy	30 (13)	15 (12)
Patient preference	10 (4)	4 (3)
Unknown	6 (3)	1 (2)
DAS28 at time of discontinuation, mean (sd)	3.19 (4.8) (n=107)	2.90 (1.4) (n=52)

All results are indicated as n (%), unless indicated otherwise. DAS: disease activity score, sd: standard deviation

LITERATURE

1. Smolen JS, Landewe RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, McInnes IB, Sepriano A, van Vollenhoven RF, de Wit M et al: EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. *Ann Rheum Dis 2020*.
2. Favalli EG, Pontikaki I, Becciolini A, Biggioggero M, Ughi N, Romano M, Crotti C, Gattinara M, Gerloni V, Marchesoni A et al: Real-life 10-year retention rate of first-line anti-TNF drugs for inflammatory arthritides in adult- and juvenile-onset populations: similarities and differences. *Clin Rheumatol 2017, 36(8):1747-1755*.
3. Favalli EG, Raimondo MG, Becciolini A, Crotti C, Biggioggero M, Caporali R: The management of first-line biologic therapy failures in rheumatoid arthritis: Current practice and future perspectives. *Autoimmun Rev* 2017, 16(12):1185-1195.
4. Kalden JR, Schulze-Koops H: Immunogenicity and loss of response to TNF inhibitors: implications for rheumatoid arthritis treatment. Nat Rev Rheumatol 2017, 13(12):707-718.
5. Hyrich KL, Watson KD, Silman AJ, Symmons DP, British Society for Rheumatology Biologics R: Predictors of response to anti-TNF-alpha therapy among patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. *Rheumatology (Oxford)* 2006, 45(12):1558-1565.
6. Hernandez MV, Sanchez-Piedra C, Garcia-Magallon B, Cuende E, Manero J, Campos-Fernandez C, Martin-Domenech R, Del Pino-Montes J, Manrique S, Castro-Villegas MC et al: Factors associated with long-term retention of treatment with golimumab in a real-world setting: an analysis of the Spanish BIOBADASER registry. *Rheumatol Int* 2019, 39(3):509-515.
7. Klareskog L, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, Martín Mola E, Pavelka K, Sany J, Settas L et al: Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. *Lancet* 2004, 363(9410):675-681.
8. Reumatoïde Artritis (RA) [https://richtlijnendatabase.nl/richtlijn/reumato_de_artritis_ra/startpagina_-_reumatoide_artritis.html]
9. Dennis G, Jr., Holweg CT, Kummerfeld SK, Choy DF, Setiadi AF, Hackney JA, Haverty PM, Gilbert H, Lin WY, Diehl L et al: Synovial phenotypes in rheumatoid arthritis correlate with response to biologic therapeutics. *Arthritis Res Ther* 2014, 16(2):R90.
10. Favalli EG, Biggioggero M, Marchesoni A, Meroni PL: Survival on treatment with second-line biologic therapy: a cohort study comparing cycling and swap strategies. *Rheumatology (Oxford)* 2014, 53(9):1664-1668.
11. Rubbert-Roth A, Szabo MZ, Kedves M, Nagy G, Atzeni F, Sarzi-Puttini P: Failure of anti-TNF treatment in patients with rheumatoid arthritis: The pros and cons of the early use of alternative biological agents. *Autoimmun Rev* 2019, 18(12):102398.
12. de Man YA, Dolhain RJ, van de Geijn FE, Willemsen SP, Hazes JM: Disease activity of rheumatoid arthritis during pregnancy: results from a nationwide prospective study. *Arthritis Rheum* 2008, 59(9):1241-1248.
13. Soliman MM, Ashcroft DM, Watson KD, Lunt M, Symmons DP, Hyrich KL, British Society for Rheumatology Biologics R: Impact of concomitant use of DMARDs

on the persistence with anti-TNF therapies in patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. *Ann Rheum Dis* 2011, 70(4):583-589.

14. Iannone F, Gremese E, Atzeni F, Biasi D, Botsios C, Cipriani P, Ferri C, Foschi V, Galeazzi M, Gerli R et al: Longterm retention of tumor necrosis factor-alpha inhibitor therapy in a large italian cohort of patients with rheumatoid arthritis from the GISEA registry: an appraisal of predictors. *J Rheumatol* 2012, 39(6):1179-1184.
15. Ogawa Y, Takahashi N, Kaneko A, Hirano Y, Kanayama Y, Yabe Y, Oguchi T, Fujibayashi T, Takagi H, Hanabayashi M et al: Association between seropositivity and discontinuation of tumor necrosis factor inhibitors due to ineffectiveness in rheumatoid arthritis. *Clin Rheumatol* 2019, 38(10):2757-2763.

CHAPTER 3

GRADUAL TAPERING TNF-INHIBITORS VERSUS CONVENTIONAL SYNTHETIC DMARDS AFTER ACHIEVING CONTROLLED DISEASE IN PATIENTS WITH RHEUMATOID ARTHRITIS; FIRST YEAR RESULTS OF THE RANDOMISED CONTROLLED TARA-STUDY

E van Mulligen, PHP de Jong, TM Kuijper, M van der Ven, CMY Appels, C Bijkerk, JB Harbers, YA de Man, ETH Molenaar, I Tchetverikov, YPM Goekoop-Ruiterman, D van Zeben, JMW Hazes, AEAM Weel, JJ Luime

Ann Rheum Dis 2019;78:746-753.

ABSTRACT

Objectives The aim of this study is to evaluate the effectiveness of two tapering strategies after achieving controlled disease in patients with rheumatoid arthritis (RA), during one year of follow-up.

Methods In this multicenter single-blinded (research nurses) randomised controlled trial RA patients were included who achieved controlled disease, defined as a DAS≤2.4 and a SJC≤1, treated with both a conventional synthetic DMARD (csDMARD) and a TNF-inhibitor. Eligible patients were randomised into gradual tapering csDMARDs or TNF-inhibitors. Medication was tapered if the RA was still under control, by cutting the dosage into half, a quarter and thereafter it was stopped. Primary outcome was proportion of patients with a disease flare, defined as DAS>2.4 and/or SJC>1. Secondary outcomes were DAS, quality of life (EQ5D) and functional ability (HAQ-DI) after one year and over time.

Results A total of 189 patients were randomly assigned to tapering csDMARDs (n=94) or tapering anti-TNF (n=95). The cumulative flare rates in the csDMARD and anti-TNF tapering group were respectively 33% (24-43%, 95% CI) and 43% (33-53%, 95% CI) (p=0.17). Mean DAS, HAQ-DI and EQ-5D did not differ between tapering groups after one year and over time.

Conclusion Up to 9 months, flare rates of tapering csDMARDs or TNF-inhibitors were similar. After one year, a non-significant difference was found of 10% favouring csDMARD tapering. Tapering TNF-inhibitors was therefore not superior to tapering csDMARDs. From a societal perspective it would be sensible to taper the TNF-inhibitor first, because of possible cost reductions and less long-term side effects.

INTRODUCTION

Treatment outcomes of rheumatoid arthritis (RA) have improved enormously during the past decades due to earlier detection of the disease, a treat-to-target approach and intensified treatment, especially combination therapy with conventional synthetic disease modifying anti-rheumatic drugs (csDMARDs) and biological DMARDs (bDMARDs). As a result, 50-60% of early RA patients are able to reach low disease activity or even sustained remission.[1-4] Because of these improved outcomes, it is nowadays more common to taper medication in RA patients, who are in sustained remission. This is in accordance with current treatment guidelines.[4] However, an optimal tapering approach, including in which order, still has to be unraveled.

The benefits of tapering treatment are (i) a decreased risk of long-term adverse events due to immunosuppression, i.e. increased infection risk and possibility of malignancy development, (ii) a reduction of health care costs, especially when biologicals are tapered, and (iii) a possibly improved compliance.[5, 6] On the other hand, tapering treatment may lead to more transient or persistent disease flares with potential harmful consequences.[1, 7, 8]

Previous studies have shown that it is possible to taper DMARDs in various ways, which has been extensively reviewed by several research groups.[7, 9-15] bDMARDs are most frequently completely withdrawn. However, with this tapering strategy the risk of disease flares in the first year of follow-up is very high. Other bDMARD tapering studies used a dose-reduction approach, which resulted in less disease flares. However, to our knowledge no randomised trials have been performed that investigate which DMARD should be tapered first.

Therefore, the aim of this study is to compare the effectiveness of two tapering strategies, namely gradually tapering csDMARDs or TNF-inhibitors, in RA patients with controlled disease under a combination of csDMARDs and a TNF-inhibitor.

PATIENTS AND METHODS

Study design

Data were used from a clinical trial (NTR2754) – namely, TApering strategies in Rheumatoid Arthritis (TARA). TARA, a multicenter, single-blinded (research nurses) randomised trial, was carried out in twelve rheumatology centers in the south-western part of the Netherlands. Inclusion started in September 2011 and ended July 2016. Medical ethics committees of each participating center approved the protocol and all patients gave written informed consent before inclusion. Patients

Adult RA patients with controlled disease, defined as a disease activity score (DAS) ≤2.4 and a swollen joint count (SJC) ≤1 at two consecutive time points within a 3-month interval, with a combination of a csDMARD and TNF-inhibitor, were included. Exclusion criteria were: (1) not being able to understand, speak and write in Dutch; (2) being diagnosed with a psychiatric or personality disorder; and (3) tapering or stopping therapy due to other reasons.

Randomisation and blinding

Patients were randomised using minimization randomisation stratified for center. Trained research nurses, blinded to the allocated treatment arm throughout the study, examined patients and calculated the DAS.

Tapering schedule

Patients were randomised into gradual tapering their csDMARD or TNF-inhibitor. csDMARD tapering was realized by cutting the dosage into half, a quarter and thereafter it was stopped. The TNF-inhibitor was tapered by doubling the dose interval, followed by cutting the dosage into half, and thereafter it was stopped. The total tapering schedule took 6 months, with dose adjustments every 3 months as long as there was still a controlled disease. At the start of the study, patients were asked to refrain from glucocorticoids (GCs). There were no restrictions on the use of NSAIDs or intra-articular GC injections.

If a disease flare occurred, defined as DAS>2.4 and/or SJC>1, tapering was stopped and the last effective treatment, when RA was under control, was restarted. In case of

a flare, one intra-muscular GC injection was allowed as bridging therapy. After a flare, no further attempts were taken to taper medication during the remainder of the first year of follow-up.

Outcomes

The primary outcome was the proportion of patients with a disease flare within one year. Secondary endpoints were disease activity, functional ability, quality of life, medication usage, and radiographic progression.

Disease activity was measured with the DAS. Functional ability was measured with the health assessment questionnaire disability index (HAQ-DI).[16] Higher HAQ-DI scores indicate poorer function. Quality of life was measured with the European Quality of Life – 5 Dimensions (EQ-5D) and short form 36 (SF36).[17-19] A higher EQ-5D index or SF36 score indicates a better quality of life. Radiographic progression was measured with the modified total Sharp score (mTSS).[20] Radiographs were scored chronologically by two out of three qualified assessors, who were blinded for study allocation and the identity of the patients.[21] Median mTSS are reported.[22] The weighted overall κ was 0.75 with >99% agreement. The percentage of patients with radiographic progression, defined as a change in mTSS>0.5 and >0.9 (the smallest detectable change), are given.[22]

Follow-up and assessments

Treatment strategies were tightly controlled, with patients being examined at baseline and every 3 months thereafter. At each time point the DAS, medication usage , development of complications and self-reported questionnaires were collected, except for hand and foot radiographs, which were obtained at baseline and after one year of follow-up.

Safety monitoring

Safety monitoring took place according to Dutch guidelines, and included laboratory tests every 3 months.[23-25] The medication was stopped or the dosage was lowered in case of adverse events related to medication use.

Statistical analysis

The TARA study was a superiority trial, powered to detect a 20% difference in flare rates between both tapering strategies. Based on related prospective cohort studies from 2011 and before, following assumptions were made: (1) 40% of the patients tapering their TNF-inhibitors to half will have controlled disease after 6 months, and (2) 60% of the csDMARD-tapering group will have controlled disease after 6 months. [26-28] Therefore, to detect this 20% difference using a significance level of α=0.05 and a power of 80%, 107 patients were needed in each treatment arm, also taking a 10% dropout ratio into account.

Outcomes were calculated in an intention-to-treat analysis, using all available data. Differences in cumulative flare rates between groups were analysed with a logistic regression model. To account for stratified randomisation by center, intercepts for each center were included. Flare-free survival was visualized with Kaplan-Meier curves. Descriptive statistics were used to assess the proportion of patients with a controlled disease after 12 months of follow-up. A linear mixed model with maximum likelihood optimisation was used to compare DAS, HAQ-DI, and EQ-5D over time. Random intercepts were included for both hospital and individual patients. Residual correlation was modeled by inclusion of an autoregressive order correlation structure. In the final model the differences in evolution over time for the outcome DAS, HAQ-DI, and EQ-5D between the two groups were assessed.

Statistical comparison of the baseline characteristics and outcomes were made by Student's t test, χ^2 test, or Wilcoxon rank-sum test, when appropriate.

All data was analyzed using STATA15. A p-value $\leq$0.05 was considered statistically significant.

RESULTS

Patients

A total of 330 patients were assessed for eligibility and 189 of those were randomly assigned to tapering their csDMARD (n=94) or tapering their TNF-inhibitor (n=95). Most patients who were not eligible did not meet the inclusion criteria for remission or refused participation (figure 1). During the first year of follow-up 14 patients withdrew from the study, mainly because of refraining from further participation (figure 1).

Table 1 shows the baseline characteristics for both tapering strategies. Patients had an average symptom duration of 6.8 years and were predominantly female (66.1%) with an average age of 56.6 years. Baseline mean (sd) HAQ-DI was 0.52 (0.47) and 0.47 (0.53) and EQ-5D was 0.86 (0.12) and 0.87 (0.11) for respectively the csDMARD and TNF-inhibitor tapering group.

At baseline, 81% of the csDMARD tapering group and 88% of the TNF-inhibitor tapering group was in remission (DAS<1.6) (table 1). The majority of patients in the csDMARD and TNF-inhibitor tapering group used MTX (respectively 97% and 86%) in combination with etanercept (respectively 54% and 55%) or adalimumab (respectively 39% and 42%). Oral glucocorticoids were taken by 4 (4%) patients in the csDMARD tapering group and 2 (2%) patients in the TNF-inhibitor tapering group, while NSAIDs were taken by 14 (15%) and 20 (21%) patients (table 1).

At baseline respectively 39% and 27% of patients within the csDMARD or TNF-inhibitor group had erosive disease.

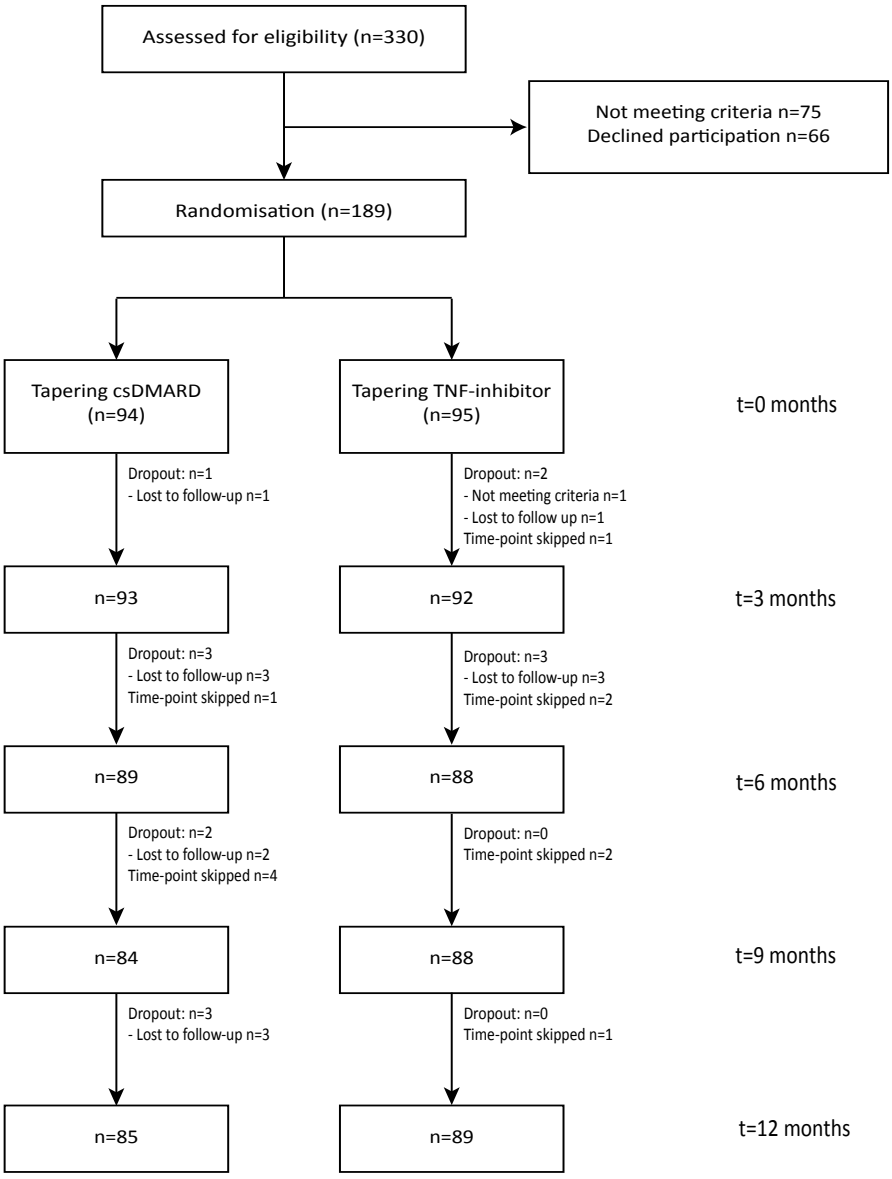


Figure 1 Trial profile and patient participation. Results are shown as number of patients. csDMARDs: conventional synthetic disease modifying anti-rheumatic drugs

Table 1 Baseline characteristics of the csDMARD tapering group and the TNF-inhibitor tapering group.

Characteristics	Tapering csDMARD (n=94)	Tapering TNF-inhibitor (n=95)
Demographic		
Age (years), mean (95% CI)	55.9 (53.0-58.8)	57.2 (55.0-59.4)
Gender, female, n (%)	67 (71)	58 (61)
Disease characteristics		
Symptom duration (years), median (IQR)	6.0 (4.1-8.5)	6.4 (4.2-8.9)
RF positive, n (%)	50 (57)	59 (65)
ACPA positive, n (%)	62 (71)	67 (75)
Disease activity		
DAS44, mean (95% CI)	1.1 (0.9-1.2)	1.0 (0.9-1.1)
DAS clinical remission, DAS44<1.6, n (%)	76 (81)	87 (88)
TJC44, median (IQR)	0 (0-2)	0 (0-1)
SJC44, median (IQR)	0 (0-0)	0 (0-0)
VAS disease activity (0-100mm), median (IQR)	20 (4-32)	12 (4-23)
ESR in mm/h, median (IQR)	8 (3-14)	8 (2-15)
CRP in mg/L, median (IQR)	2.2 (1-5)	2 (1-6)
Use of csDMARDs*		
MTX, n (%)	90 (96)	84 (88)
SASP, n (%)	10 (11)	12 (13)
HCQ, n (%)	24 (26)	37 (39)
Leflunomide, n (%)	2 (2)	4 (4)
Use of TNF-inhibitor		
Etanercept, n (%)	51 (54)	52 (55)
Adalimumab, n (%)	37 (39)	40 (42)
Others, n (%) **	6 (6)	3 (3)
Radiographs (hand/foot)		
mTSS (0-488), median (IQR)	2 (0-6.5)	1 (0-3.5)
Erosion score (0-280), median (IQR)	0 (0-2.5)	0 (0-2)
JSN score (0-168), median (IQR)	0.5 (0-2.5)	0 (0-2.5)
Erosive disease, n(%) ***	37 (39)	26 (27)
Patient-reported outcomes		
HAQ-DI, mean (95% CI)	0.52 (0.42-0.62)	0.47 (0.35-0.58)
SF-36, median (IQR)		
- PCS	43 (29-48)	47 (39-51)
- MCS	60 (56-63)	57 (51-62)
EQ5D index, mean (95% CI)	0.86 (0.83-0.88)	0.87 (0.85-0.89)

*some patients used a combination of csDMARDs, **certolizumab or golimumab, *** Erosive disease is characterized as having >1 erosion in three separate joints. ACPA: anti-citrullinated protein antibody; CI: confidence interval; CRP: C-reactive protein; csDMARD: conventional

synthetic disease modifying anti-rheumatic drug; DAS44: disease activity score measured in 44 joints; ESR: erythrocyte sedimentation rate; EQ5D: European Quality of Life – 5 Dimensions; HAQ-DI: Health Assessment Questionnaire Disability Index; HCQ: hydroxychloroquine; IQR: interquartile range; JSN: joint space narrowing; MCS: mental component summary; mTSS: modified Sharp/Van der Heijde score; MTX: methotrexate; PCS: physical component summary; RF: rheumatoid factor; SASP: sulfasalazine; SF-36: short form 36; SJC: swollen joint count; TJC: tender joint count; VAS: visual analogue scale

Outcomes

After one year of follow-up, the cumulative flare rate was 33% (24-43%, 95% CI) in the csDMARD and 43% (33-53%, 95% CI) in the anti-TNF tapering group (figure 2). This means that 63/94 (67%) in the csDMARDs tapering group and 54/95 (57%) in the TNF-inhibitor tapering group still had a well-controlled RA (p=0.17). Of the patients who flared and restarted the last effective treatment strategy, 46% regained a DAS<2.4 within 3 months, which increased to 67% by 6 months. Two patients (1%) were unable to get back in remission within the first year.

No significant differences were seen in DAS (p=0.72), HAQ-DI (p=0.63), and EQ-5D (p=0.58) after one year between both tapering strategies (table 2). Also over time, the DAS (p=0.49), and EQ-5D (p=0.35) were not significantly different between both tapering strategies (figure 3). Although the TNF-inhibitor tapering group seems to have lower HAQ-DI scores over time, this was not significantly different (p=0.15)(figure 3). Over time, the patients with a disease flare increased and thus the proportion of patients with a DAS<2.4 decreased in both tapering strategies. A similar trend was seen for the HAQ-DI and EQ-5D over time (figure 3).

Median mTSS scores were 2 (IQR 0-6.5) in the csDMARD and 1 (IQR 0-4) in the TNF-inhibitor tapering group after one year of follow-up (table 2). Radiographic progression was seen in 5% of the csDMARD tapering group and 6% of the anti-TNF tapering group (p=0.82). Also, the cumulative probability plots were overlapping (figure 3B).

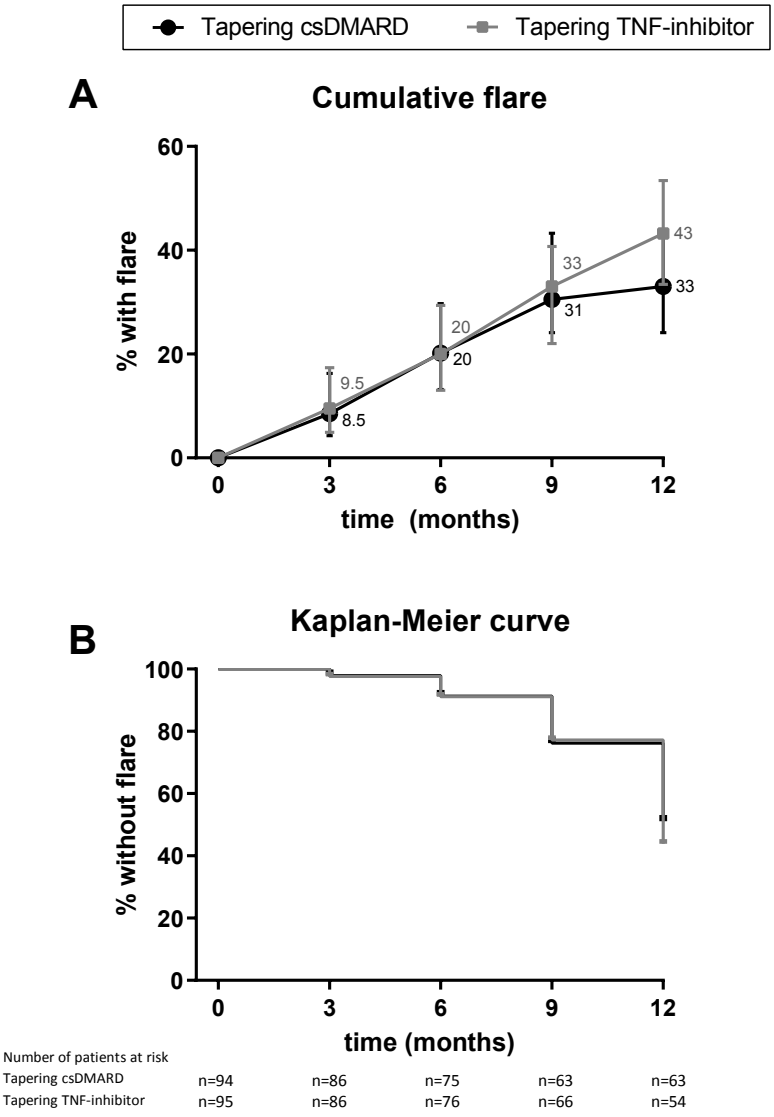


Figure 2 Percentages flares and Kaplan-Meier curves for maintenance of controlled disease in the first 12 months. % with flare indicates the cumulative number of patients with flares. Error bars indicate 95% confidence intervals. Kaplan-Meier curves indicate loss of controlled disease (DAS44>2.4 and/or SJC>1) over time. Numbers below the Kaplan-Meier curve indicate the number of patients at risk per time point. csDMARD: conventional synthetic disease modifying anti-rheumatic drug.

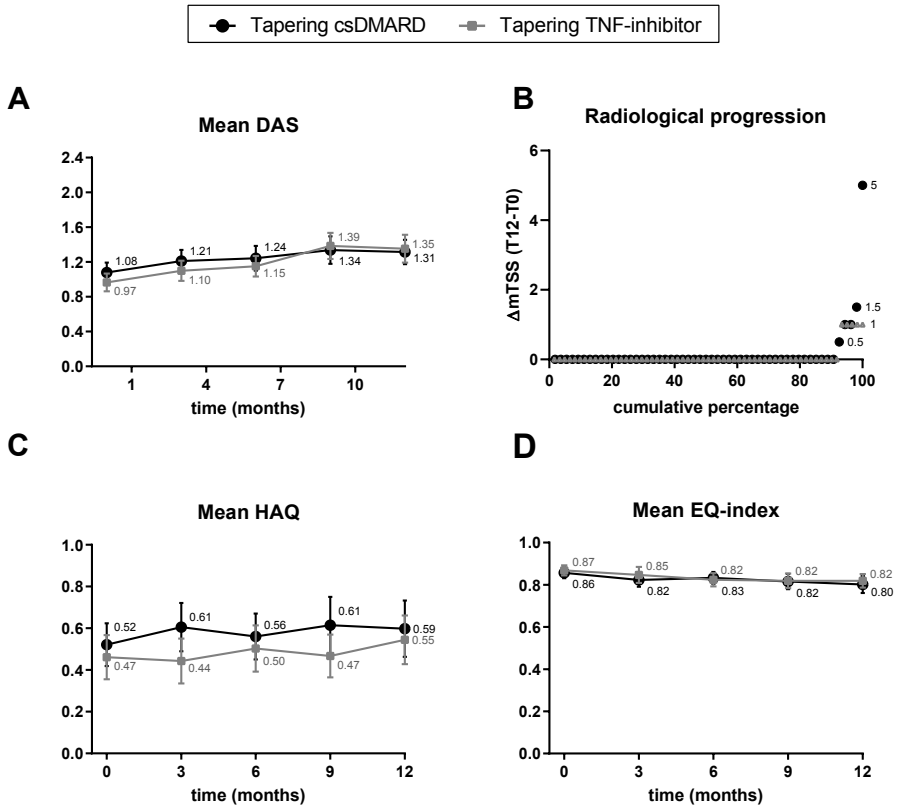


Figure 3 Disease activity, cumulative probability plot for radiological progression, functional ability and quality of life over time per tapering arm. (A, C, D) Error bars indicate 95% confidence intervals. (B) Each point represents radiological progression (T12-T0) of an individual patient, measured with the modified Sharp/Van der Heijde (mTSS) score at 0 and 12 months. csDMARD: conventional synthetic disease modifying anti-rheumatic drug; DAS: disease activity score; EQ-5D index: Dutch EuroQol index; HAQ: Health Assessment Questionnaire; mTSS: modified Sharp/Van der Heijde score.

Table 2 Clinical response after 12 months for both tapering groups, according to intention-to-treat.

Clinical response after 12 months	Tapering csDMARD (n=85)	Tapering TNF -inhibitor (n=89)
Disease activity		
DAS44, mean (95% CI)	1.31 (1.17-1.46)	1.35 (1.19-1.51)
TJC44, median (IQR)	0 (0-2)	0 (0-3)
SJC44, median (IQR)	0 (0-0)	0 (0-1)
VAS disease activity (0-100mm), median (IQR)	17 (5-36)	19 (6-42)
ESR in mm/h, median (IQR)	11 (5-21)	11 (4-19)
CRP in mg/L, median (IQR)	2.9 (1-6)	4 (1-9)

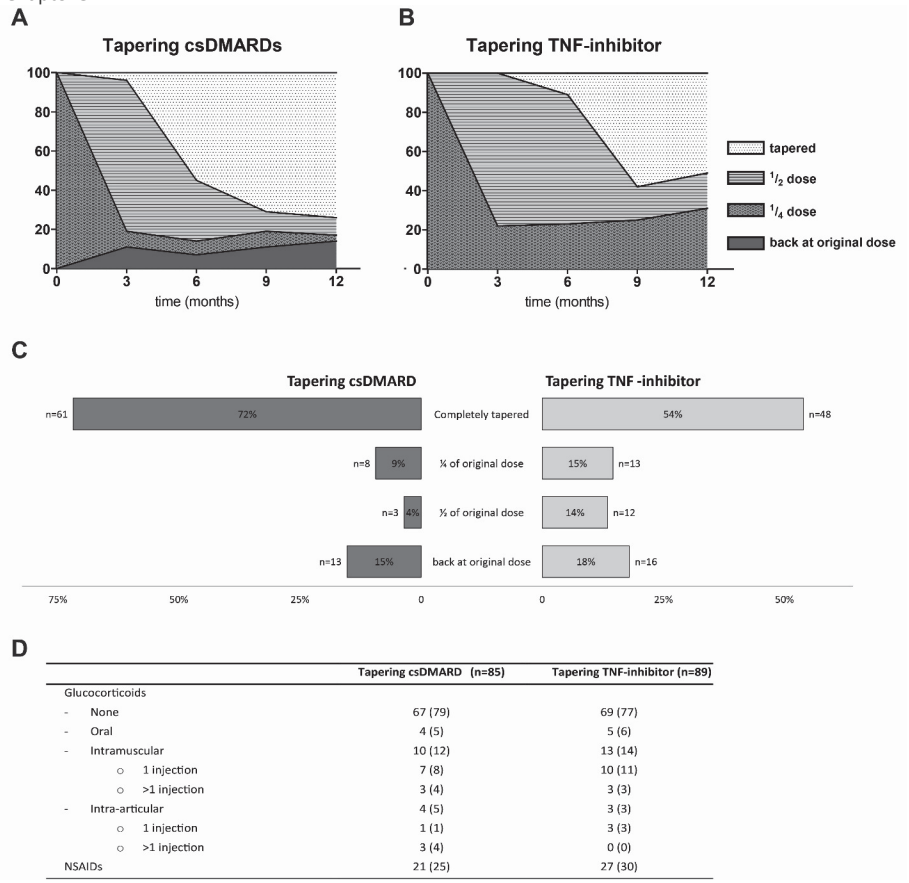
Table 2 (Continued)

Clinical response after 12 months	Tapering csDMARD (n=85)	Tapering TNF -inhibitor (n=89)
DAS clinical remission, DAS44<1.6, n (%)	57 (69)	58 (66)
ΔDAS44 (T12-T0), mean (95% CI)	0.28 (0.16-0.40)	0.40 (0.22-0.57)
Radiographic progression (hand/foot)		
mTSS (0-488), median (IQR)	2 (0-6.5)	1 (0-4)
Erosion score (0-280), median (IQR)	0.5 (0-2)	0 (0-2)
JSN score (0-168), median (IQR)	0.5 (0-2.5)	0 (0-2.5)
ΔmTSS (T12-T0), median (IQR)	0 (0-0)	0 (0-0)
Patients with progression >0.5, n (%)	4 (5)	5 (6)
Patients with progression >0.9, n (%)	4 (5)	5 (6)
Erosive disease, n(%)*	37 (44)	30 (34)
Patient-reported outcomes		
HAQ-DI, mean (95% CI)	0.59 (0.46-0.73)	0.55 (0.43-0.66)
ΔHAQ-DI (T12-T0), mean (95% CI)	0.05 (-0.05-0.13)	0.07 (-0.01-0.16)
SF-36, median (IQR)		
- PCS	43 (32-50)	44 (35-50)
- MCS	58 (53-62)	59 (51-62)
EQ5D index, mean (95% CI)	0.80 (0.76-0.84)	0.82 (0.79-0.85)
ΔEQ5D index (T12-T0), mean (95% CI)	-0.06 (-0.09- -0.02)	-0.05 (-0.08- -0.02)

*Erosive disease is characterized as having >1 erosion in three separate joints. CI: confidence interval; CRP: C-reactive protein; csDMARD: conventional synthetic disease modifying anti-rheumatic drug; DAS44: disease activity score measured in 44 joints; ESR: erythrocyte sedimentation rate; EQ5D: European Quality of Life – 5 Dimensions; HAQ-DI: Health Assessment Questionnaire Disability Index; IQR: interquartile range; JSN; Joint space narrowing, MCS: mental component summary; mTSS: modified Sharp/Van der Heijde score; PCS: physical component summary; SJC: swollen joint count; TJC: tender joint count; VAS: visual analogue scale

Treatment

After 12 months, 58 patients in the csDMARD tapering group and 45 patients in the TNF-inhibitor tapering group completely tapered their medication (p=0.09). On the other hand, 8 and 16 patients were using the same dosage as at start of the trial. The remaining patients were able to taper their medication partially (figure 4C). The course of the tapering schedule is visualized in figure 4A and 4B. There was an overall significant difference in tapering status after 12 months of follow-up between the two tapering strategies (p=0.02). During the follow-up period we found no significant differences in glucocorticoid and NSAID usage between both tapering groups (figure 4D).



D

	Tapering csDMARD (n=85)	Tapering TNF-inhibitor (n=89)
Glucocorticoids		
- None	67 (79)	69 (77)
- Oral	4 (5)	5 (6)
- Intramuscular	10 (12)	13 (14)
○ 1 injection	7 (8)	10 (11)
○ >1 injection	3 (4)	3 (3)
- Intra-articular	4 (5)	3 (3)
○ 1 injection	1 (1)	3 (3)
○ >1 injection	3 (4)	0 (0)
NSAIDs	21 (25)	27 (30)

Figure 4 Status of tapering in the first year of follow-up. (A, B) Overview of tapering status per time point. Results are shown as percentages of patients. According to protocol, the doses were halved every 3 months, starting at T0, and after 6 months patients could stop their tapered medication when they were still in a controlled disease state. (C) Tapering status after 12 months. Columns indicate the percentage of patients that tapered medication until the indicated amount of the original dose. (D) Overview of glucocorticoids and NSAID use in the first year of follow-up. csDMARD: conventional synthetic disease modifying anti-rheumatic drug; NSAID: Non-Steroidal Anti-Inflammatory Drugs.

Adverse events

In the csDMARD tapering group 82 adverse events were self-reported versus 98 in the TNF-inhibitor tapering group (online supplemental table S1). Serious adverse events (SAEs) were seen in 10 (12%) patients tapering csDMARDs and 5 (6%) patients tapering TNF-inhibitor (p=0.3, online supplemental table S1). Reported SAEs were hospitalization, herpes zoster infection, basal cell carcinoma, large cell lung carcinoma, and a bruised rib. None of the SAEs were considered to be related to the trial treatment.

DISCUSSION

Tapering csDMARDs resulted in a 33% (24-43%, 95% CI) flare rate (DAS44>2.4 and/or a SJC>1), while tapering TNF-inhibitors gave a 43% (33-53%, (95% CI) flare rate over a one-year period in the randomised controlled TARA trial. At 12 months, 103 (59%) patients were able to stop either their TNF-inhibitor or csDMARD, while 47 (37%) patients were using a lower dosage. Clinical and patient reported outcomes were comparable in both tapering groups over time and after one year of follow-up. Also, no significant differences in adverse events or radiological progression were seen between both tapering strategies.

Nowadays, more RA patients achieve a state of sustained remission, which makes them eligible for tapering treatment. This is reflected in current EULAR recommendations for the management for RA. The advice is to taper DMARD therapy in RA patients who are in sustained remission in the following ordering: glucocorticoids, bDMARDS, and csDMARDs.[4] Our results and the fact that TNF-blockers are more expensive than csDMARDs support aformentioned tapering order.

The majority of previous tapering trials focused on the withdrawal of TNF-inhibitors alone. Flare rates for tapering TNF-inhibitors varied between 51% and 77%. The POET study, for example, reported a 51.2% flare rate (DAS28>3.2 or ΔDAS28>0.6) after stopping the TNF-inhibitor.[7] The STRASS showed a 76.6% flare rate (DAS28>2.6 or ΔDAS28>0.6) when extending the dosage interval of the TNF-inhibitor.[14] The DRESS study reported a 55% flare rate (ΔDAS28-CRP>0.6) after a dose-reduction of the TNF-inbitor.[13] Finally, the PRESERVE trial reported a 57.4% flare rate (DAS28>3.2) when the TNF-inhibitor was stopped, and a 20.9% flare rate (DAS28>3.2) when the TNF-inhibitor dose was cut into half.[15] Only few randomised controlled trials investigated tapering of csDMARDs, but the majority looked at the combined tapering of csDMARDs and biologicals. Flare rates within these studies varied between 35 – 56%.[1, 6, 29-32]

Although flare rates of aforementioned studies are similar to or higher than our findings, direct comparison is difficult, because of the differences in study design. The most important study design differences are: (1) no common definition for relapse or flare, (2) no comparison between tapering of csDMARDs and TNF-inhibitors, and (3) DMARD therapy could only be tapered or stopped once during follow-up. If we would use other criteria to define a flare in the TARA population, we would observe higher hypothetical flare rates. We would have encountered a 74.1% flare rate if using

DAS28>3.2 or ΔDAS28>0.6, an 80.5% flare rate if using DAS28>2.6 or ΔDAS28>0.6, a 52.3% flare rate if we use ΔDAS28-CRP>0.6, and a 39.1% flare rate if using DAS28>3.2. Mostly, these flare rates are higher than our reported flare rates, but are similar to previous mentioned trials. This indicates that our criteria were more strict than other studies, but that flare rates are comparible between the tapering studies.

Also, the flare duration was longer in the TARA trial compared to other trials, which could be due to the measurement intervals of 3 months. If patients did not have a controlled disease 3 months after flare, we assumed that the duration of flare was 6 months. That might be a reason that our results seem to have a long flare duration compared to the DOSERA or DRESS study, in which they knew the exact duration of flare in weeks.[12, 13]

In this study there are several strengths and limitations. Strengths of the study are that we performed a randomised controlled trial to asess tapering in RA patients with a controlled disease. The TARA trial is one of the first trials which assess differences in tapering strategies, and elaborates on current viewpoints concerning tapering treatment, instead of only determining if tapering is feasible or not.

Some limitations should be noted as well. First of all, inclusion was terminated earlier due to difficulties with recruiting. This was due to the initial inclusion criteria being too strict (DAS≤1.6), and the start of another trial (POET study) which used the same pool of eligible patients.[7] The study sample size was based on a 20% difference between both tapering strategies resulting in 96 patients per arm. We found, however, a 10% difference with 85 patients in the csDMARD and 89 patients in the TNF-inhibitor arm. This resulted in a power of 70% instead of 80%. For this reason, we performed a worst case scenario analysis to see if our results were valid. We used the following assumptions: (1) all extra included patients in the csDMARD tapering group had no flares and (2) all extra included patients in the TNF-inhibitor tapering group flared. This analysis showed an 18% difference in flares, which is still below the 20% difference on which our power calculation was based. Therefore, we think our current results and conclusions are valid.

Second, rheumatologist could have only referred patients who achieved low disease activity quickly and had less severe disease and, therefore, creating selection bias. However, we think that our target population is the same as the one we would apply our results to, because those are the patients who are suitable for tapering and are willing to taper their medication. Second, only research nurses, who did the DAS

assessment, were blinded. Rheumatologists, therefore, knew the tapering strategy of their patients. This design was chosen to mimic daily practise as much as possible. However, it could be a possible source of bias, since rheumatologist might prefer one of the two tapering strategies and would possibly treat patients differently depending on the tapering strategy.

Third, the time frame of follow-up was only one year. Although the differences in flare rates were not significantly different between both tapering strategies, the largest difference was seen at 12 months. Data of the second year are needed to investigate if this difference will increase.

Last, we encountered 19% protocol violations , which could underestimate the effect of one of the two tapering strategies. We analyzed the type of violations and we can conclude that most protocol violations were randomly distributed over the two treatment arms and were made due to a treat-to-target approach.

To ensure optimal rheumatic care in the future, efficient use of biological treatment is needed.[33] By tapering medication, costs can be reduced, especially when tapering bDMARDs. On the other hand, 38% of the patients in the TARA study flared within the first year, which may have a direct impact on patients' lives (i.e. worker productivity and unemployment). Therefore, it is important to know which tapering strategy is most cost-effective, which will be addressed in a follow-up analysis.

In conclusion, the TARA study showed that up to 9 months, flare rates of tapering csDMARDs or TNF-inhibitors were similar. After one year, a non-significant difference in flare rates was found of 10% in favor of csDMARD tapering. Tapering TNF-inhibitors was therefore not superior to tapering csDMARDs. From a societal perspective it would be sensible to taper the TNF-inhibitor first, because of possible cost reductions and less long-term side effects.

Supplemental table S1 Adverse Events

	Tapering csDMARD (n=85)	Tapering TNF-inhibitor (n=89)
Serious Adverse events (SAEs)	10 (12)	5 (6)
Hospitalization	7 (8)	4 (4)
- Total hip or knee replacement surgery	3 (4)	1 (1)
- Pneumonia	1 (1)	-
- Decompression shoulder	1 (1)	-
- Pancreatitis	1 (1)	-
- Angina pectoris	1 (1)	-
- Peripheral vascular disease	-	1 (1)
- Myocardial infarction	-	1 (1)
Herpes zoster	1 (1)	-
Basal cell carcinoma	1 (1)	-
Large-cell lung carcinoma	1 (1)	-
Bruised rib	-	1 (1)
Observed AEs		
Patients >1 AE	23 (27)	21 (24)
Gastrointestinal complaints	20 (24)	27 (31)
Fatigue	7 (8)	11 (13)
Off day	10 (12)	5 (6)
Hair loss	3 (4)	4 (5)
Acne	0 (0)	3 (3)
Mouth	1 (1)	5 (6)
Headache	2 (2)	0 (0)
Skin irritation	18 (21)	16 (18)
Pain of injection	16 (19)	12 (14)
Fear of injection	5 (6)	2 (2)

All results are shown as n(%). AEs: adverse events. csDMARD: conventional synthetic disease modifying anti-rheumatic drug.

Supplemental table S2 Detailed overview of percentages of patients in the indicated tapering status per time point

	T0	T3	T6	T9	T12
Tapering csDMARD					
Full dose	0%	11%	7%	11%	10%
½ dose	100%	8%	7%	8%	15%
¼ dose	0%	77%	31%	10%	7%
Tapered	0%	4%	55%	71%	68%
Tapering TNF-inhibitor					
Full dose	0%	4%	11%	15%	18%
½ dose	100%	18%	12%	10%	11%
¼ dose	0%	78%	66%	17%	20%
Tapered	0%	0%	11%	58%	51%

csDMARD: conventional synthetic disease modifying anti-rheumatic drug

LITERATURE

1. Kuijper TM, Luime JJ, de Jong PH, Gerards AH, van Zeben D, Tchetverikov I, de Sonnaville PB, van Krugten MV, Grillet BA, Hazes JM *et al*: Tapering conventional synthetic DMARDs in patients with early arthritis in sustained remission: 2-year follow-up of the tREACH trial. *Ann Rheum Dis* 2016, 75(12):2119-2123.
2. Scott DL, Wolfe F, Huizinga TWJ: Rheumatoid arthritis. *Lancet* 2010, 376:15.
3. Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC, Vaysbrot E, McNaughton C, Osani M, Shmerling RH *et al*: 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol* 2016, 68(1):1-26.
4. Smolen JS, Landewe R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, Nam J, Ramiro S, Voshaar M, van Vollenhoven R *et al*: EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017, 76(6):960-977.
5. Fautrel B, den Broeder AA: De-intensifying treatment in established rheumatoid arthritis (RA): Why, how, when and in whom can DMARDs be tapered? *Best Pract Res Clin Rheumatol* 2015, 29(4):550-565.
6. Haschka J, Englbrecht M, Hueber AJ, Manger B, Kleyer A, Reiser M, Finzel S, Tony H-P, Kleinert S, Feuchtenberger 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* 2016, 75(1):45-51.
7. Ghiti Moghadam M, Vonkeman HE, Ten Klooster PM, Tekstra J, van Schaardenburg D, Starmans-Kool M, Brouwer E, Bos R, Lems WF, Colin EM *et al*: Stopping Tumor Necrosis Factor Inhibitor Treatment in Patients With Established Rheumatoid Arthritis in Remission or With Stable Low Disease Activity: A Pragmatic Multicenter, Open-Label Randomized Controlled Trial. *Arthritis Rheumatol* 2016, 68(8):1810-1817.
8. Schett G, Emery P, Tanaka Y, Burmester G, Pisetsky DS, Naredo E, Fautrel B, van Vollenhoven R: Tapering biologic and conventional DMARD therapy in rheumatoid arthritis: current evidence and future directions. *Ann Rheum Dis* 2016, 75(8):1428-1437.
9. Kuijper TM, Lamers-Karnebeek FB, Jacobs JW, Hazes JM, Luime JJ: Flare Rate in Patients with Rheumatoid Arthritis in Low Disease Activity or Remission When Tapering or Stopping Synthetic or Biologic DMARD: A Systematic Review. *J Rheumatol* 2015, 42(11):2012-2022.
10. Edwards CJ, Fautrel B, Schulze-Koops H, Huizinga TWJ, Kruger K: Dosing down with biologic therapies: a systematic review and clinicians' perspective. *Rheumatology (Oxford)* 2017, 56(11):1847-1856.
11. van Herwaarden N, den Broeder AA, Jacobs W, van der Maas A, Bijlsma JW, van Vollenhoven RF, van den Bemt BJ: Down-titration and discontinuation strategies of tumor necrosis factor-blocking agents for rheumatoid arthritis in patients with low disease activity. *Cochrane Database Syst Rev* 2014(9):CD010455.
12. van Vollenhoven RF, Ostergaard M, Leirisalo-Repo M, Uhlig T, Jansson M, Larsson E, Brock F, Franck-Larsson K: Full dose, reduced dose or discontinuation of etanercept in rheumatoid arthritis. *Ann Rheum Dis* 2016, 75(1):52-58.
13. van Herwaarden N, van der Maas A, Minten MJ, van den Hoogen FH, Kievit W, van Vollenhoven RF, Bijlsma JW, van den Bemt BJ, den Broeder AA: Disease activity guided dose reduction and withdrawal of adalimumab or etanercept compared with usual care in rheumatoid

arthritis: open label, randomised controlled, non-inferiority trial. *Bmj* 2015, 350:h1389.

14. Fautrel B, Pham T, Alfaiate T, Gandjbakhch F, Foltz V, Morel J, Dernis E, Gaudin P, Brocq O, Solau-Gervais E *et al*: Step-down strategy of spacing TNF-blocker injections for established rheumatoid arthritis in remission: results of the multicentre non-inferiority randomised open-label controlled trial (STRASS: Spacing of TNF-blocker injections in Rheumatoid ArthritiS Study). *Ann Rheum Dis* 2016, 75(1):59-67.
15. Smolen JS, Nash P, Durez P, Hall S, Ilivanova E, Irazoque-Palazuelos F, Miranda P, Park MC, Pavelka K, Pedersen R *et al*: Maintenance, reduction, or withdrawal of etanercept after treatment with etanercept and methotrexate in patients with moderate rheumatoid arthritis (PRESERVE): a randomised controlled trial. *Lancet* 2013, 381(9870):918-929.
16. Siegert CE, Vleming LJ, Vandenbroucke JP, Cats A: Measurement of disability in Dutch rheumatoid arthritis patients. *Clin Rheumatol* 1984, 3(3):305-309.
17. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A: Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *Br J Rheumatol* 1997, 36(5):551-559.
18. McHorney CA, Ware JE, Jr., Raczek AE: The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993, 31(3):247-263.
19. Ware JE, Jr., Sherbourne CD: The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992, 30(6):473-483.
20. van der Heijde D: How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol* 2000, 27(1):261-263.
21. van Tuyl LH, van der Heijde D, Knol DL, Boers M: Chronological reading of radiographs in rheumatoid arthritis increases efficiency and does not lead to bias. *Ann Rheum Dis* 2014, 73(2):391-395.
22. van der Heijde D, Simon L, Smolen J, Strand V, Sharp J, Boers M, Breedveld F, Weisman M, Weinblatt M, Rau R *et al*: How to report radiographic data in randomized clinical trials in rheumatoid arthritis: guidelines from a roundtable discussion. *Arthritis Rheum* 2002, 47(2):215-218.
23. NVR Richtlijnen medicijnen: biologicals [https://www.nvr.nl/wp-content/uploads/2014/11/NVR-Medicijnen-Update_Biologicals_richtlijn-23-6-2014.pdf]
24. NVR Richtlijnen medicijnen: sulfazalazine [https://www.nvr.nl/wp-content/uploads/2014/11/NVR-Medicijnen-richtlijn-Sulfasalazine-2002.pdf]
25. NVR Richtlijnen medicijnen: methotrexaat [https://www.nvr.nl/wp-content/uploads/2014/11/NVR-Medicijnen-MTX-richtlijn-2009-update-2011.pdf]
26. Tanaka Y, Takeuchi T, Mimori T, Saito K, Nawata M, Kameda H, Nojima T, Miyasaka N, Koike T, investigators RRRs: Discontinuation of infliximab after attaining low disease activity in patients with rheumatoid arthritis: RRR (remission induction by Remicade in RA) study. *Ann Rheum Dis* 2010, 69(7):1286-1291.
27. Saleem B, Brown AK, Keen H, Nizam S, Freeston J, Karim Z, Quinn M, Wakefield R, Hensor E, Conaghan PG *et al*: Disease remission state in patients treated with the combination of tumor necrosis factor blockade and methotrexate or with disease-modifying antirheumatic drugs: A clinical and imaging comparative study. *Arthritis Rheum* 2009, 60(7):1915-1922.
28. Lagana B, Picchianti Diamanti A, Ferlito C, Germano V, Migliore A, Cremona A, Argento G, David V, Salemi S, D'Amelio R: Imaging progression despite clinical remission in early rheumatoid arthritis patients after etanercept interruption.

Int J Immunopathol Pharmacol 2009, 22(2):447-454.

29. ten Wolde S, Breedveld FC, Hermans J, Vandenbroucke JP, van de Laar MA, Markusse HM, Janssen M, van den Brink HR, Dijkmans BA: Randomised placebo-controlled study of stopping second-line drugs in rheumatoid arthritis. *Lancet* 1996, 347(8998):347-352.
30. Gotzsche PC, Hansen M, Stoltenberg M, Svendsen A, Beier J, Faarvang KL, Wangel M, Rydgren L, Halberg P, Juncker P *et al*: Randomized, placebo controlled trial of withdrawal of slow-acting antirheumatic drugs and of observer bias in rheumatoid arthritis. *Scand J Rheumatol* 1996, 25(4):194-199.
31. Klarenbeek NB, van der Kooij SM, Guler-Yuksel M, van Groenendael JH, Han KH, Kerstens PJ, Huizinga TW, Dijkmans BA, Allaart CF: Discontinuing treatment in patients with rheumatoid arthritis in sustained clinical remission: exploratory analyses from the BeSt study. *Ann Rheum Dis* 2011, 70(2):315-319.
32. Allaart CF, Lems WF, Huizinga TW: The BeSt way of withdrawing biologic agents. *Clin Exp Rheumatol* 2013, 31(4 Suppl 78):S14-18.
33. Schoels M, Wong J, Scott DL, Zink A, Richards P, Landewe R, Smolen JS, Aletaha D: Economic aspects of treatment options in rheumatoid arthritis: a systematic literature review informing the EULAR recommendations for the management of rheumatoid arthritis. *Ann Rheum Dis* 2010, 69(6):995-1003.

CHAPTER 4

RESPONSE TO THE LETTER: "TARA STUDY – A NEW PERSPECTIVE ON TAPERING DRUGS IN RA"

E van Mulligen, JMW Hazes, AEAM Weel, PHP de Jong

Ann Rheum Dis 2020;79:e80.

We are pleased about the interest in our article by Mishra *et al.* and we would like to respond to their questions so that there can be no ambiguity.[1, 2]

First of all, there is some clarification needed on the csDMARD that were used in combination with the TNF-inhibitors at baseline. In table 1 we elaborate on the different combination of csDMARDs that were used for each intervention arm separately. In the csDMARD tapering group the methotrexate (MTX) was tapered, except for the 3 patients that did not use MTX. These patients gradually tapered leflunomide (n=1) and sulfasalazine (n=2).

Table 1 Use of csDMARDs at baseline in the TARA study specified for two groups: tapering csDMARDs and tapering TNF-inhibitors.

Use of csDMARDs at baseline	Tapering csDMARD (n=93)	Tapering TNF inhibitor (n=95)
MTX monotherapy, n(%)	64 (69)	49 (52)
MTX + hydroxychloroquine, n(%)	17 (18)	27 (29)
MTX + sulfasalazine + hydroxychloroquine, n(%)	5 (5)	6 (6)
MTX + sulfasalazine, n(%)	3 (3)	2 (2)
MTX + leflunomide, n(%)	1 (1)	0 (0)
Sulfasalazine monotherapy, n(%)	0 (0)	3 (3)
Sulfasalazine + hydroxychloroquine, n(%)	2 (2)	0 (0)
Sulfasalazine + leflunomide, n(%)	0 (0)	1 (1)
Leflunomide monotherapy, n(%)	1 (1)	3 (3)
Leflunomide + hydroxychloroquine, n(%)	0 (0)	1 (1)
Hydroxychloroquine monotherapy, n(%)	0 (0)	3 (3)

csDMARD: conventional synthetic disease modifying anti-rheumatic drugs; MTX: methotrexate

Mishra *et al.* also had a question about our intention-to-treat (ITT) analysis. In an ITT analysis patients are analyzed in the groups to which they were randomised, regardless of whether they received or adhered to the allocated intervention. Therefore, in the clinical response table (table 2) of the original article we should have given the total numbers instead of the patients who were still participating in the TARA trial at 12 months.[1] If we had given the total numbers the results would be similar.

Third question was about explaining the difference between the number of patients who are in remission after 12 months of follow-up and the number of patients below the Kaplan Meier (KM)-curve at 12 months. In a KM-curve only the patients at risk are given. Patients are censored if they experience a flare or drop out, which results in a decreasing number of patients at risk over time. In table 2, on the other hand,

the number of patients in clinical remission (defined as a DAS<1.6) at 12 months of follow-up are given. Thus, the interpretation of the numbers given in the KM-curve and table 2 are different and, therefore, the numbers are non-identical.

Finally, it would be interesting to know if the primary outcome would change if we use a modified per-protocol approach as brought up by Mishra *et al.* For this reason we excluded the patients that used oral glucocorticoids, n=4 and n=5 in respectively the csDMARD and TNF-inhibitor tapering group, or had more than one intramuscular injection, n=3 in each tapering group. With aforementioned approach a 30% (95% confidence interval (CI) 21%-41%) flare rate was seen in the csDMARD tapering group, and a 39% (95% CI, 31%-52%) flare rate in the TNF-inhibitor group (p=0.15). The difference in flare rates between both tapering arms is similar to the original article.[1]

LITERATURE

1. van Mulligen E, de Jong PHP, Kuijper TM, van der Ven M, Appels C, Bijkerk C, Harbers JB, de Man Y, Molenaar THE, Tchetverikov I *et al:* Gradual tapering TNF inhibitors versus conventional synthetic DMARDs after achieving controlled disease in patients with rheumatoid arthritis: first-year results of the randomised controlled TARA study. *Ann Rheum Dis* 2019, 78(6):746-753.
2. Mishra D, Chattopadhyay A, Jain S, Dhir V: TARA Study - a new perspective on tapering drugs in RA. *Ann Rheum Dis* 2019.

CHAPTER 5

DMARD-FREE REMISSION AS NOVEL TREATMENT TARGET IN RHEUMATOID ARTHRITIS: A SYSTEMATIC LITERATURE REVIEW OF ACHIEVABILITY AND SUSTAINABILITY

E van Mulligen, M Verstappen, PHP de Jong, AHM van der Helm- van Mil

RMD Open. 2020; 6(1):e001220.

ABSTRACT

OBJECTIVES Although current treatment guidelines for rheumatoid arthritis (RA) suggest tapering DMARDs, it is unclear whether DMARD-free remission (DFR) is an achievable and sustainable outcome. Therefore we systematically reviewed the literature to determine prevalence and sustainability of DFR, and evaluated potential predictors for DFR.

METHODS A systematic literature search was performed in March 2019 in multiple databases. All clinical trials and observational studies reporting on discontinuation of DMARDs in RA-patients in remission were included. Our quality assessment included a general assessment and assessment of the description of DFR. Prevalence of DFR and its sustainability, flares during tapering and after DMARD-stop were summarized. Also, potential predictors for achieving DFR were reviewed.

RESULTS From 631 articles, 51 were included, comprising 14 clinical trials and 5 observational studies. DFR-definition differed, especially for the duration of DMARD-free state. Considering only high and moderate-quality studies, DFR was achieved in 5.0%-24.3%, and sustained DFR (duration>12 months) in 11.6%-19.4% (both relative to number of patients eligible for tapering). Flares occurred frequently during DMARD-tapering (41.8%-75.0%) and in the first year after achieving DFR (10.4%-11.8%), whilst late flares, >1 year after DMARD-stop, were infrequent (0.3%-3.5%). Many patient characteristics lacked association with DFR. Absence of auto-antibodies and shared epitope alleles increased the chance of achieving DFR.

CONCLUSIONS DFR is achievable in RA, and is sustainable in ~10%-20% of patients. DFR can become an important outcome measure for clinical trials, and requires consistency in the definition. Considering the high rate of flares in the first year after DMARD-stop, a DMARD-free follow-up of >12 months is advisable to evaluate sustainability.

INTRODUCTION

In rheumatoid arthritis (RA) early treatment, with disease-modifying anti-rheumatic drugs (DMARDs), aiming at sustained remission, is nowadays the key element of each management approach.[1, 2] As a result, RA has become a controllable disease in which sustained clinical remission is achievable for an increasing number of patients, and tapering and discontinuation of DMARDs has become of emerging interest.[3] Current international guidelines recommend tapering of DMARDs in RA patients with sustained remission.[1, 2] Nevertheless, these guidelines are less clear whether DMARDs can be stopped and the systematic literature review supportive of the most recent EULAR guidelines was not focussed on DMARD-cessation.[4]

Despite the recommendations in the guidelines, tapering of DMARDs has not been adopted structurally in many clinical practices, presumably because the risk of a disease flare[5], and because the ability to achieve and sustain DMARD-free remission (DFR) is often considered unlikely.[6] On the other hand, there is increasing interest in achieving DFR, because this is currently the best proxy for cure.[7, 8] Clinical trials occasionally report on DFR, but usually not as primary outcome. Absence of knowledge of DFR prevalence, its sustainability and the characteristics of patients achieving DFR currently hampers the use of DFR as primary outcome.[9]

We aimed to expand the comprehension of the ability to achieve and sustain DFR in RA. Therefore, we conducted a systematic literature search. In addition to the DFR prevalence and sustainability, potential predictors for achieving DFR were explored.

METHODS

Search strategy and selection criteria

This systematic literature review was conducted in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines and the Cochrane review handbook.[10, 11] The protocol was registered in the International Prospective Register of Systematic Reviews (CRD42019132558).[12]

The search strategy was developed and performed in collaboration with an experienced librarian (JS). Key terms used for the search were 'Rheumatoid arthritis', 'Antirheumatic drugs', 'Discontinuation' and 'Remission'. These search items were

translated into multiple matching synonyms in order to broaden our results. All search elements were combined with the Boolean operators AND/OR. PubMed, Embase, Web of Science, COCHRANE Library, Emcare and Academic Search Premier were systematically searched (supplementary table S1).

All observational cohorts and clinical trials reporting on discontinuation of DMARDs in RA-patients, in remission, were included. Study selection was independently carried out by two reviewers (MV and EvM). Cases of disagreement were discussed until consensus was reached. First, all obtained titles were screened, subsequently abstracts were reviewed after which full-text articles were screened for the predefined in- and exclusion criteria (supplementary table S2). If multiple articles were based on the same study, the article which described prevalence and sustainability of DFR most clearly was selected. Subsequently, the article describing the longest follow-up was used for data extraction.

Data extraction

A standardized data collection form was used to extract the following information: study design, patient characteristics, interventions, glucocorticoid (GCs) usage, organization of follow-up, outcome measures and loss to follow-up (LTFU) (supplementary table S3). Furthermore, data regarding eligibility criteria for tapering, tapering methods, numbers of patients tapering, description and timing of achieving DFR, sustainment of DFR over time and the occurrence of flares were extensively explored. Also information regarding predictors of DFR was collected. Data extraction was done independently by two reviewers (MV and EvM), disagreements were discussed until consensus was reached.

If the methods were incomplete or unclear, the methods of the original study could be used if a reference was available. Clinical trials and observational studies were handled separately, because of fundamental differences in the study design, which could influence achievement and sustainment of DFR; i.e. protocolized versus non-protocolized tapering, frequency of monitoring and duration of follow-up.

Quality assessment

Our study-quality assessment consisted of 2 parts, namely a general assessment and an assessment of the description of DFR. For the general quality assessment

we used 13 predefined quality criteria, which were based on Cochrane guidelines (supplementary table S4).[11] The general study-quality was considered 'good' if >75% (≥10 items) of these criteria were scored positive. For the DFR quality assessment we used the following criteria: (1)'DFR definition', referring to whether a definition (e.g. remission criterion) of DFR was included, and (2)'DFR-duration', referring to whether information on the time between DMARD-stop and being appointed as DFR (i.e. the duration of DMARD-free status) was reported. Specific emphasis was put on duration of DMARD-free state since this attains insight in sustainability of DFR. When both DFR-quality criteria were scored positive, DFR-quality was regarded as 'good'.

Studies were regarded as 'high-quality' if the general, as well as the DFR quality were good . When general study-quality was good but only one DFR-criterion was fulfilled, studies were regarded as 'moderate-quality'. Studies lacking both DFR-criteria, or without a good general quality assessment were scored as 'low-quality'.

Data analysis

Extracted data was used to calculate DFR prevalence, defined as the proportion of patients achieving DFR, compared to those eligible for tapering medication. For each prevalence the confidence interval (CI) was calculated. Patients were considered eligible for tapering when they had achieved remission and subsequently were allowed to start tapering their medication. GCs were also considered as DMARDs. We specifically chose not to use the total study population as denominator, because in some studies specific groups of patients were not allowed to taper their medication due to study protocol.

Sustained DFR (SDFR) was defined as the percentage of patients with a DFR-duration of >12 months since DMARD-stop, relative to the number of patients eligible for tapering. Reported flares were categorized and summarized according to the time-period in which they occurred: (i) during tapering, (ii) in the first year after achieving DFR ('early flares') and (iii) after more than one year of DFR ('late flares'). Results on DFR were summarized in a narrative overview, also in relation to study-quality. Due to expected heterogeneity in study design and study populations, pooled effect estimates were not calculated.

Additionally, the data was reviewed on potential predictors for achieving DFR. We used the same methods for data extraction and assessment as described for DFR-

prevalence. Predictors of DFR were summarized. Results on variables evaluated in more than one high or moderate-quality article were graphically presented, based on statistical significance obtained with regression analysis. If univariate and multivariate analysis were both conducted, results of the multivariate analysis were used. For each predictor, the number of studies and the total number of patients within these studies were presented and the direction of the effect was indicated.

RESULTS

Study selection

Our search resulted in 631 articles, of which 51 articles were considered eligible for inclusion (figure 1). These 51 articles comprised data from 19 studies, 14 clinical trials and 5 observational cohorts.

Quality assessment

Both the quality of the study in general and the description of DFR were evaluated, resulting in a final quality rating. Eleven out of 14 clinical trials and 2 out of 5 observational cohorts showed a good general quality (table 1). Notably, the tapering methods were better described for clinical trials than for observational cohorts. Of the 13 studies with a good general quality, 7 fulfilled both quality criteria for DFR and were regarded as high-quality. These 7 high-quality studies comprised 5 clinical trials and 2 observational cohorts. Of the remaining 6 studies, two studies were of moderate quality since only one DFR criterion was fulfilled. The four other studies did not fulfil any DFR-quality criteria and were regarded low quality (table 1).

Because of fundamental differences in study design, DFR-prevalence and flare rates from clinical trials and observational cohorts were presented separately. Also, only high or moderate-quality studies were presented in the result section. Nonetheless, all prevalence, including those of low-quality studies, can be found in table 2.

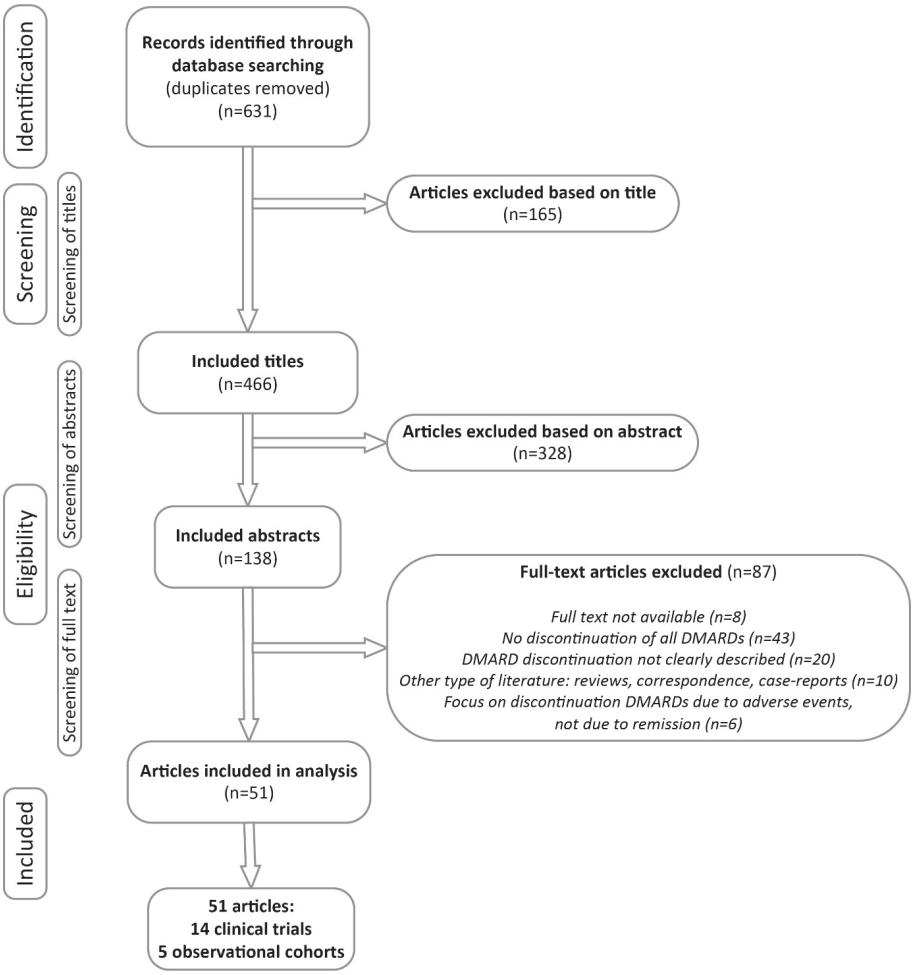


Figure 1 Flow diagram of study selection.

DMARDs: disease-modifying antirheumatic drugs.

Table 1 Assessment of general study quality and DFR quality, resulting in final categorization as high, moderate or low-quality study.

		Clinical trials														Observational studies				
		BeSt	IMPROVED	AVERT	tREACH	U-ACT-early	ACT-RAY	El Miedany et al.	PRIZE	RETRO	Ten Wolde et al.	SUPRISE	Brocq et al.	Kita et al.	DREAM trial	Leiden EAC	DREAM cohort	Tiippanna-Kinnunen et	ESPOIR	ERAS
DMARD-free remission	DFR definition *(description of DFR-criteria)*	+	+	+	+	+	+	-	-	-	-	-	-	-	-	+	+	-	+	+
	DFR-duration *(description of DFR period)*	+	+	+	+	+	-	+	-	-	-	-	+	+	-	+	+	-	+	+
DFR-quality		✓	✓	✓	✓	✓	±	±					±	±		✓	✓		✓	✓
Study population	Selection of patients *(description in-/exclusion criteria)*	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Criteria for RA diagnosis	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+
	Baseline characteristics *(description of characteristics)*	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Randomization *for different study treatments*		+	+	+	+	+	+	+	+	+	+	+	-	-	-	*n.a.*				
Blinding *of study treatment*		±	±	+	±	+	+	-	+	-	+	-	-	-	-	*n.a.*				
Intervention	Treatment strategies *(description of strategies)*	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
	Cut-off point tapering *(description of cut-off point)*	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	-	-
	Tapering method *(description of methods)*	+	+	+	+	+	+	+	+	+	+	-	+	-	-	-	-	-	-	-
Follow-up	Organization follow-up *(frequency of monitoring)*	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	-
	Lost-to-follow up *(description of LTFU)*	+	+	+	+	+	+	?	?	?	-	+	+	+	-	-	+	+	+	+
Data analysis & presentation	Outcome reporting	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	?	+
	Analysis techniques *(description of techniques)*	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	?	?
	Missing data *(handling of missing data)*	+	+	?	?	+	?	-	±	?	?	?	?	-	?	+	?	?	?	?
General study quality		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				✓	✓			
Combined quality		HQ	HQ	HQ	HQ	HQ	MQ	MQ	LQ	LQ	LQ	LQ	LQ	LQ	LQ	HQ	HQ	LQ	LQ	LQ

Studies were assessed for quality of DMARD-free remission: whether definition (yes "+" or no "–") and duration of drug-free state were reported (yes "+" or no "–"). DFR-quality was considered good ("✓ ") when both items were scored as '+', and moderate ('±') when only 1 of 2 was scored as good. Subsequently, studies were assessed on general study-quality. Criteria for general study-quality could be scored: "+" indicating sufficient, "-" indicating not sufficient, "±" indicating moderate, "?" indicating unclear reporting and quality could not be assessed. Study-quality was considered good ("✓ ") when minimally 75% (10 items) were scored as '+'. DFR: DMARD-free remission; DMARD: disease-modifying anti-rheumatic drug; LTFU: lost to follow-up; n.a.: not applicable; RA: rheumatoid arthritis. The combined study-quality was considered high ('HQ') when both DFR-quality, as study-quality was good. It was considered moderate ('MQ') when DFR-quality was moderate, and study-quality was good. Low quality ('LQ') indicates studies with either insufficient DFR-quality and/or study-quality.

Table 2 Prevalence of DMARD-free remission and flares.

Clinical trials														
						Tapering			DMARD-free remission					
Study	**Inclusion period**	**N**	**Study pop charact.[c]**	**Treatment/Intervention**	**FU**	**Tapering criteria**	**N**	**Flares during tapering (definition flare)**	**% DFR achieved* (≤12m)**	**Time in DFR (months)**	**Early flares (≤12m)**	**% Sustained DFR** (>12m)**	**Late flares (>12m)**
									Definition of DFR				
BeSt[13]	2000 - 2008	508	Early RA 1987 ACR crit HDA at BL	1: Monotherapy (126) 2: Step-up combi (121) 3: Initial combi (133) 4: Combi with IFX (128) *After 2y tapering possible.*	60m (5y)	Tapering: DAS44<2.4 DMARD stop: DAS44<1.6 min 6m	-	-	**22.6%**[B] (115/508) DAS44 <1.6	5	**10.4%** 53/508 DAS44 ≥1.6	**11.6%**[B] (59/508) at 5y FU DAS44 <1.6 min 12m	-
IMPROVED[14]	2007 - 2010	479	Early RA 2010 ACR crit HDA at BL	0- 4m MTx + Pred. >4-8m DAS28<1.6: taper DAS28>1.6: 1: Triple csDMARDs 2: ADA+MTx	60m (5y)	DAS44<1.6	-	-	**23.0%** (110/479) DAS44 <1.6 or Boolean	12	-	**19.4%** (93/479) at 5y FU DAS44 <1.6 or Boolean 12m	**3.5%** (17/479) at 5y FU
AVERT[15]	2010 - 2014	351	Early RA Diagnosis by exp. opinion HDA at BL ACPA+	0-52w 1: ABA+MTx (119) 2: ABA (116) 3: MTx (166) >52w Withdrawal	18m	DAS28-CRP<3.2	223	-	**18.4%** (41/223) DAS28 <2.6	5	-	-	-
tREACH[16]	2007 - 2011	281	Early RA 2010 ACR crit HDA at BL	1: Triple therapy (183) 2: MTx (98) *Tapering at any time in FU.*	24m	DAS44<1.6 min 2 visit	141	**41.8%** 59/141 (DAS44 ≥2.4)	**5.0%** (7/141) DAS44 <1.6	6	-	-	-

Table 2 Clinical trials (Continued)

Clinical trials														
						Tapering			DMARD-free remission					
Study	**Inclusion period**	**N**	**Study pop charact.[c]**	**Treatment/Intervention**	**FU**	**Tapering criteria**	**N**	**Flares during tapering (definition flare)**	**% DFR achieved* (≤12m)**	**Time in DFR (months)**	**Early flares (≤12m)**	**% Sustained DFR** (>12m)**	**Late flares (>12m)**
									Definition of DFR				
U-Act-Early[17]	2010 - 2012	317	Early RA 1987/2010 ACR crit HDA at BL	1: TCZ+MTx (106) 2: MTx (103) 3: TCZ (108) *Tapering at any time in FU*	24m	DAS28<2.6 SJC<4 min 24w	-	-	**24.3%**[B] (77/317) DAS28<2.6 & SJC≤4	3	-	-	-
ACT-RAY[20]	2009 - 2013	556	Establ. RA 1987 ACR crit HDA at BL	0-52w TOCI +MTx >52w *or* TOCI (279/277) T2T + tapering (472)	12-36m	DAS28<2.6 min 12w	472	**42.4%** 200/472 (Expert opinion)	**5.9%**[B] (28/472) DAS28 <2.6	single timepoint	-	-	-
El Miedany et al[26]	-	157	RA duration n.r. 2010 ACR crit Remission at BL	Arm 1-3: Taper DMARDs Arm 4: Stop all DMARDs Arm 5: Control	12m	DAS28<2.6 min 6m	32	**75.0%**[B] 24/32 (DAS>3.2)	**21.9%** (7/32) DAS28 <2.6	12	-	-	-
PRIZE[18]	2009 - 2012	306	Early RA 1987 ACR crit HDA at BL	0-52w ETA+MTx 52-91w ETA+MTx *or* MTx *or* placebo 91-117w Withdrawal	29m	DAS28<3.2	132	-	**46.9%** (62/132) DAS28 <2.6	22-24w	-	-	-
RETRO[22]	2010 - 2013	101	Establ. RA 2010 ACR crit. Remission at BL	1: Continue (38) 2: Tapering (36) 3: Stop DMARD (27)(6m 50%)	12m	DAS28<2.6 min 6m	27	**51.9%** (14/27) (DAS28>2.6)	**48.1%** (13/27) DAS28 <2.6	6	-	-	-

Table 2 Clinical trials (Continued)

Clinical trials														
						Tapering			DMARD-free remission					
Study	**Inclusion period**	**N**	**Study pop charact.[c]**	**Treatment/Intervention**	**FU**	**Tapering criteria**	**N**	**Flares during tapering (definition flare)**	**% DFR achieved* (≤12m)**	**Time in DFR (months)**	**Early flares (≤12m)**	**% Sustained DFR** (>12m)**	**Late flares (>12m)**
									Definition of DFR				
Ten Wolde et al.[23]	-	285	Establ. RA 1987 ACR crit Remission at BL	1: Continue (142) 2: Switch placebo (143)	12m	ARA remission (5/6 crit)	143	**37.1%**[B] (53/143) (Synovitis)	**58%** (83/143) ARA remission (5/6)	12	-	-	-
Brocq et al.[24]	1995 - 2005	21	Establ. RA 2010 ACR crit Remission at BL	TNFi treatment at inclusion Intervention: abrupt stop TNFi	12m	DAS28<2.6 min 6m	7	**57.1%**[B] (4/7) (DAS28 >3.2)	**28.6%**[B] (2/7) DAS28 <2.6	12	-	-	-
SURPRISE[21]	2009 - 2012	233	Establ. RA 1987 ACR crit HDA at BL	0-52w TCZ+MTx or TCZ >52w Stop TCZ	24m	DAS28<2.6	53	**66%**[B] (35/53) (n.r.)	**26.4%**[B] (14/53) DAS28 <2.6	12	-	-	-
Kita et al.[19]	2008 - 2009	13	Early RA 2010 ACR crit HDA at BL ACPA+	0-52w Treat-to-target >52w Stop all DMARDs	24m	SDAI & BME-33% on MRI	5	**20%**[B] (1/5) (n.r.)	**60%**[B] (3/5) SDAI remission	12	-	-	-
DREAM trial[25]	2008 - 2010	187	Establ. RA 1987 ACR crit LDA at BL	Tapering after 4y TCZ mono-therapy	12m	DAS28<3.2	187	**72.5%**[B] (136/187) (DAS28>3.2)	**9.1%**[B] (17/187) DAS28 <2.6	12	-	-	-

Table 2 Clinical trials (Continued)

Observational studies														
						Tapering			DMARD-free remission					
Study	**Inclusion period**	**N**	**Study pop charact.[c]**	**Treatment/Intervention**	**FU**	**Tapering criteria**	**N**	**Flares during tapering**	**% DFR achieved* (≤12m)**	**Time in DFR (months)**	**Early flares (≤12m)**	**% Sustained DFR** (>12m)**	**Late flares (>12m)**
									Definition of DFR				
Leiden EAC[28]	1993 - 2011	889	Early RA 1987 ACR crit. HDA at BL	'93-'95 '96-'98 '99-'04 >2005 NSAIDs Mild DMARDs Initial MTx DAS-steered	1-18y	-	-	-	-	12	-	**17.8%** (158/889) after 1-18y FU No synovitis min 12m	**0.3 %** (3/889)
DREAM cohort [31]	2006 - 2009	229	Early RA expert opinion (79% 1987) HDA at BL	Treat-to-target, steered at DAS28<2.6: Initial MTx monotherapy, if DAS28>2.6 + SSZ if DAS28>3.2 TNF-inhibitor	5y	DAS28<2.6 min 6m	-	-	**23.6%**[B] (54/229) DAS28 <2.6	6	**11.8%** (27/229) DAS28>2.6	**11.8%**[B] (27/229) after 5y FU DAS28<2.6 min 6m	-
Tiippana-Kinnunen et al.[29]	1986 - 1989	70	Early RA 1958/1987 crit. HDA at BL	Sawtooth strategy	15y	Clinical remission or minor disease activity	-	-	**28.6%**[B] (20/70) ARA remission[E]	-	**12.9%** (9/70)	**15.7%**[B] (11/70) after 15y FU ARA remission	-
ESPOIR[32] D	2000 - 2005	533	Early RA Clin diagnosis HDA at BL	Treated with cDMARDs	5y	-	-	-	-	12	-	**5.4%** (29/533) No synovitis min 12m	-

Table 2 Clinical trials (Continued)

Observational studies													
						Tapering			DMARD-free remission				
Study	Inclusion period	N	Study pop charact.[c]	Treatment/Intervention	FU	Tapering criteria	N	Flares during tapering	% DFR achieved* (≤12m)	Time in DFR (months)	Early flares (≤12m)	% Sustained DFR** (>12m)	Late flares (>12m)
									Definition of DFR				
ERAS[30] [D]	1986 - 1996	895	Early RA 1987 ACR crit. HDA at BL	Rheumatologist preference, predominantly MTx, SSZ, HCQ	10y	-	-	-	-	12	-	**9.4%** (84/895) No synovitis min 12m	-

High-quality studies are indicated in dark grey, moderate-quality studies are indicated in light grey, and low-quality studies are indicated in white.*Percentage of patients who achieved DFR divided by patients eligible for tapering. **Percentage of patients who sustained DFR for more than 12 months divided by patients eligible for tapering.

† Potential use of intra-articular or systemic corticosteroids, or use of GCS was not clearly described due to which use was doubtful

[A] DMARDs were discontinued abruptly without gradual tapering method. [B] Clinical remission defined as no tender joints , no swollen joints, no joint pain by history, ESR<30(female/<20(male) for minimal 12 months. Or prolonged symptom-free phase of disease with minor disease activity. [c] Longstanding RA was defined as a disease duration of more than 2 years. All shorter disease- and symptom durations were classified as early RA. In the supplementary table (S1) specific duration of disease and symptom duration can be found. [D] Only minimal information could be extracted from the articles in which this study was mentioned. Therefore information is missing, which is not due to insufficient quality of the article. [E] ARA remission: morning stiffness absent (or not exceeding 15 minutes), no fatigue, no joint pain by history, no joint tenderness, no joint or tendon sheath swelling, no elevation of ESR (in 5/6, fatigue is not included in the criteria).

ACR: American College of Rheumatology, ADA: Adalimumab, BL: Baseline, bDMARDs: biological DMARDs, crit: criteria, csDMARDs: conventional DMARDs, DAS: Disease activity score, DFR: DMARD-free remission, Establ.: Established, ETA: Etanercept, FU: Follow-Up, HCQ: Hydroxychloroquine, HDA: High disease activity, IFX: Infliximab, LDA: Low disease activity, rem: remission, MTx: Methotrexate, SSZ: Sulfasalazine, n.r.: not reported, TCZ: Tocilizumab, TNFi: TNF-α inhibitor

Clinical trials

Study characteristics

Study populations varied in RA classification (1987 versus 2010 criteria), disease stage/duration (early versus established) and disease activity (supplementary table S5). Overall, trials were performed in two 'settings': early, DMARD-naïve RA and established RA. Studies including early RA had a treat-to-target approach and when remission was achieved, DMARDs were tapered. This was all conducted in a relative short period of time (n=7)[13-19]. The established RA studies (disease duration 3.1-11.3 years, n=6) either included patients with active disease who first changed DMARD-treatment and subsequently became eligible for tapering (n=2)[20, 21], or selected patients who were in longstanding remission and were directly considered eligible for tapering (n=4)[22-25]. All established RA studies were of low-quality, except 1 which was of moderate-quality.[20] One study, including patients in sustained remission, did not report disease duration.[26]

DMARD tapering

Tapering of DMARDs was initiated when patients fulfilled the study-specific eligibility criteria for tapering, in which some were stricter than others (supplementary table S5). Methods of tapering varied from immediate DMARD-stop to one-by-one gradual tapering of DMARDs over the course of a year. In general, tapering of biologic DMARDs took place before tapering of conventional synthetic DMARDs. Flare rates during tapering ranged from 41.8%-75.0% (table 2, figure 2).

Definitions of DMARD-free remission

Overall, the remission criterion used to define DFR was mainly DAS44 or DAS28 remission. The DFR rates were either given as a point-prevalence, thus at the moment of DMARD-stop, or combined with a minimal DFR-period of several months (table 2, figure 2). Nevertheless, most studies did not put much emphasize on a minimal duration of drug-free state as requirement to achieve DFR. Importantly, 3 studies that clearly defined DFR (2 high-quality, 1 moderate-quality) allowed i.a. or oral GCs during DFR, without reporting the actual use.[13, 17, 20]

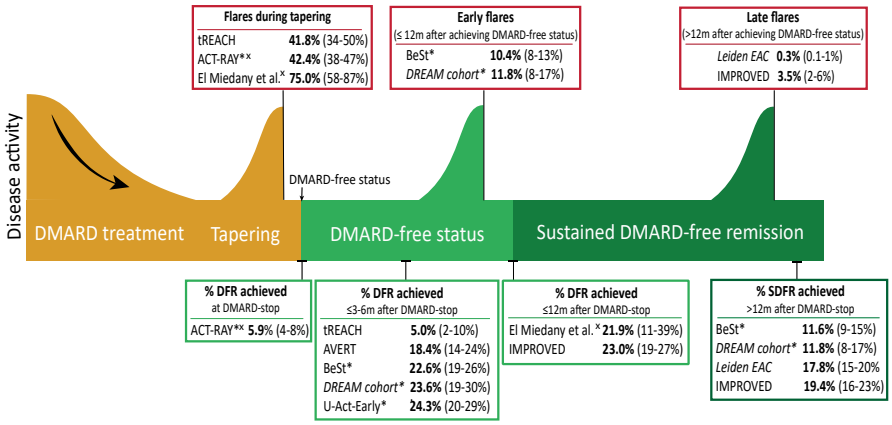


Figure 2 Summary of flare rate and DFR-prevalence, all as percentage of the number of patients that were eligible for DMARD-tapering, depicted on a timeline. DFR-prevalence was grouped by the duration of DFR.

Legend: Data is presented as DFR percentage (confidence interval). Data were based on high or moderate-quality studies. Prevalence and confidence intervals were calculated using number of DFR patients divided by the number of patients eligible for tapering. Results from observational studies are indicated in italic.

*indicates that studies that allowed the use of i.a. or systemic corticosteroids in patients that were considered to be in DFR (absolute number of patients that used corticosteroids after DMARD-stop was not reported). x indicates moderate-quality studies.

DMARD: Disease-modifying antirheumatic drugs, DFR: DMARD-free remission, SDFR: sustained DMARD-free remission.

Prevalence of DMARD-free remission

In the 5 high-quality clinical trials, the reported prevalence of DFR (DFR <12 months) ranged from 5.0% to 24.3% (relative to the number of patients eligible for tapering). The 2 moderate-quality studies reported a DFR-prevalence of 5.9% respectively 21.9% (table 2, figure 2). When studies allowed GCs while being in DFR were excluded, DFR occurred in 5.0%-23.0%. SDFR (DFR >12 months) was only reported in 2 clinical trials, and showed a prevalence of 11.6% and 19.4% (relative to patients eligible for tapering).

Evaluation of DFR-prevalence, in high and moderate-quality studies, in relation to the trial 'settings' was hampered by the fact that only 1 study was performed in established RA where DMARDs were tapered after prolonged remission[20], revealing a DFR-prevalence of 5.9% compared to the prevalence of 5.0%-24.3% in studies that tapered DMARDs in early RA.[13-17, 27]

Early flares (≤12 months after DMARD-stop) were reported in one high-quality study and occurred in 10.4% of patients eligible for tapering. Late flares (>12 months after DMARD-stop) were reported by another study and occurred in 3.5% of patients (table 2).

Observational cohorts

Study characteristics

Patients included in the observational cohorts were diagnosed between 1986-2011 (n=5). Patients in the observational cohorts were, compared to clinical trials, included in an earlier time period, but had a longer follow-up. Diagnosis was based on the 1987-criteria,[28-30] or expert opinion.[31, 32] Treatment was less protocolized compared to the clinical trials and a treat-to-target approach was only used in three studies,[28, 29, 31] of which two had a high quality (table 2).

DMARD tapering

Eligibility for tapering was only clearly reported in 1 study.[31]

Definitions of DMARD-free remission

Remission within DFR was defined as the absence of clinical synovitis (table 2), except for 1 study that used a DAS28 cutoff (DAS28<2.6).[31] All 5 observational cohorts reported on SDFR (DFR>12 months), whereas 1 also reported on DFR after 6 months. In two studies, of which 1 high-quality study, i.a. and oral glucocorticoid were allowed while being in DFR; the actual use was not reported.

Prevalence of DMARD-free remission

DFR prevalence (<12 months) was 23.6% of patients eligible for tapering, and was reported in one high-quality study.[31] The prevalence of SDFR ranged from 11.8%-17.8% (relative to patients eligible for tapering)(table 2, figure 2)[28, 31] If we exclude the studies that allowed GCs during DFR, one high-quality study remained with a SDFR-prevalence of 17.8%.[28] We did not compare DFR-prevalence between studies that did and did not apply a treat-to-target approach, because all studies without a treat-to-target approach were of low-quality.

Early flares (≤12 months after DMARD-stop) were reported in one high-quality study and occurred in 11.8% of patients eligible for tapering. Late flares were reported by the other high-quality study and were seen in 0.3% of patients eligible for tapering (table 2).

Predictors of DFR

All factors that were analyzed for their potential association with achieving DFR were evaluated (supplementary table S6). Due to heterogeneity in evaluated effect estimates, effect sizes could not be compared and meta-analyses not performed. For predictors that were studied in more than one high or moderate-quality study the association with achieving DFR was summarized in figure 3 (see also supplementary table S7). The figure includes information on the number of studies with/without an association, the total number of patients in these studies, and the directionality of the effect (if present). The absence of autoantibodies and HLA shared epitope alleles were predictive for achieving DFR. Many patient characteristics (e.g. age, BMI, SJC, ESR, erosions at baseline) were not associated with the chance of achieving DFR. For some characteristics findings were inconsistent. Results on symptom duration, for example, showed ambiguous results (supplementary table S6/7).

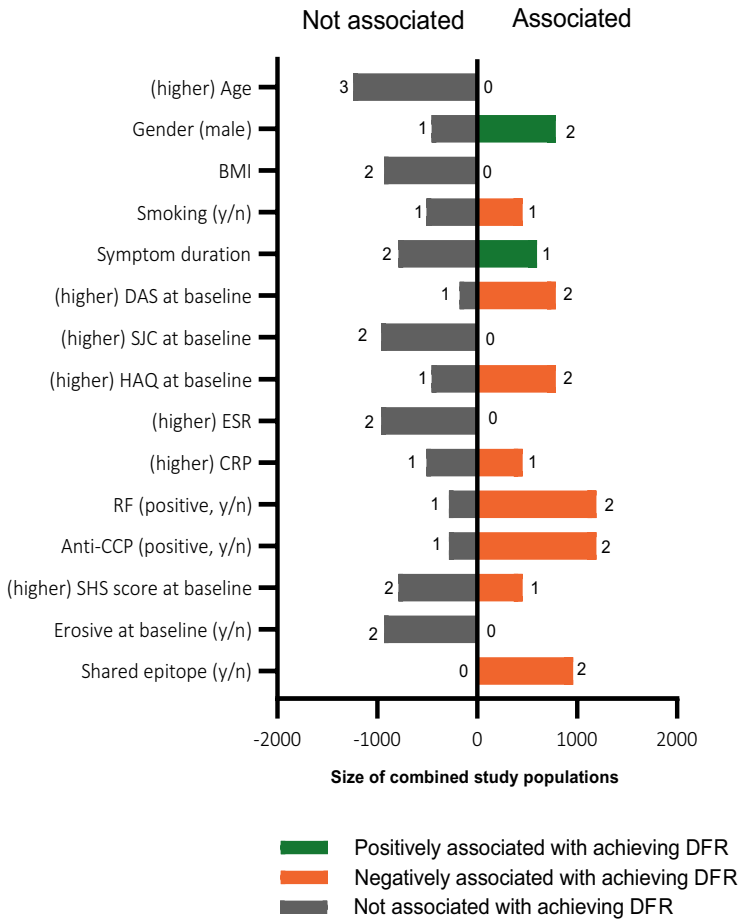


Figure 3 Overview of studied predictors of achieving DMARD-free remission. Data are presented from variables that were reported in >1 study, based on statistical significance obtained in regression analysis. If both univariable and multivariable regression was applied, the result of the multivariable regression was used. Presented are the absence (left panel) and presence of an association with achieving DFR over time (right panel), the number of studies are indicated per predictor, the total number of patients in these studies is plotted on the x-axis. The directionality of the effect is indicated in colours, green indicates an increased risk of achieving DFR, red indicates a decreased risk of achieving DFR. For symptom duration no differentiation was made for analyses using this as continuous or categorical variable. BMI: Body Mass Index, DAS: Disease Activity Score, SJC: Swollen Joint Count, HAQ: Health Assessment Questionnaire, ESR: Estimated Sedimentation Rate, CRP: C-Reactive Protein, RF: Rheumatoid Factor

DISCUSSION

This systematic literature review was conducted in accordance with PRISMA guidelines and provides insight in the occurrence and sustainability of DFR in RA. The prevalence of DFR (DFR ≤12 months) was 5.0%-24.3%[14-17, 27], while SDFR (DFR >12 months) was achievable in 11.6%-19.4% of patients eligible for tapering.

Remission criteria used to define DFR varied widely, and the temporal aspect (sustainability) varied as well or was not reported. Moreover, in some studies concomitant use of glucocorticoids was allowed while patients were in DFR. This might falsely inflate DFR-prevalence, but to what extend this occurred is unclear as actual use was not reported. Exclusion of aforementioned studies did not affect our results. To increase homogeneity, quality criteria were used, and final conclusions were only based on high and moderate-quality studies, which resulted in a narrative overview of DFR prevalence (figure 2).

We observed different DFR prevalence depending on the duration of the DFR period. To allow fair comparison of DFR-prevalence we categorized the duration of DFR in groups. SDFR was defined as a DMARD-free period >12 months. Higher prevalence were observed when DFR had a less stringent criterion for sustainability (figure 2). In line with this, flares occurred most often during tapering and in the first months after DMARD-stop. This time-effect underlines the relevance of defining sustainability of DFR in future studies.

DFR and SDFR might be fundamentally different. Short-term DFR might indicate that disease activity was suppressed, but not necessarily resolved, and could revive after disappearance of suppressive treatment. Moreover, early flares (≤12 months after DFR) occur more often than late flares (>12 months after DFR), which might indicate that auto-immunity was not completely silenced. In our opinion, patients in SDFR (DFR >12 months) better resemblance silencing of autoimmunity and may have achieved a proxy for cure. Therefore, SDFR can become an important outcome for clinical trials. Because, late flares (often occurring years after DMARD-stop) might be pathophysiological different from early flares; it is an interesting subject for future studies to explore the triggers or pathophysiologic mechanisms involved in late reactivation of the auto-immune process.

Notably, despite differences in study design, the DFR-prevalence observed in observational cohorts and clinical trials were comparable. This supports the

robustness of the observed frequencies. We were unable to investigate how long remission should be sustained before tapering can be initiated, because too few high-quality studies were performed in patients with established RA and longstanding remission. Additionally, due to an insufficient amount of studies, nothing can be said about the change of achieving DFR after treatment with certain conventional or biologic DMARDs.

We could not evaluate whether the method of tapering influenced the frequency of SDFR. It has been suggested that gradual tapering results in less flares compared to abrupt cessation.[2] Also the stringency of the remission criterion for initiation of tapering might be of influence, whereby less stringent criteria might increase the risk of flares. Evaluation of the methods of DMARD-tapering was beyond the scope of this review, and a relevant subject for further studies as insight in the most effective tapering method may positively influence the chance to achieve SDFR. Another issue for further studies is the assessment of the likelihood to achieve remission for patients that flare after having been in DFR. From studies on patients that flared during tapering it is known that the majority of patients achieve remission by restarting the same DMARD.[33] Whether this is similar for patients that flare after DMARD-stop is not yet systematically studied.

Studying the prevalence of DFR and predictors for DFR does not answer the question whether the absence of clinical signs and symptoms without treatment exhibited the natural course of RA in these patients[34], or was induced by DMARD-treatment. This could not be answered within our SLR, nor could we compare studies for treatment intensity (e.g. reflected by treat-to-target) due to the lack of high-quality studies without a treat-to-target approach. One high-quality study compared a treat-to-target approach that aimed for a DAS<1.6 with an approach that aimed for a DAS<2.4, and reported that patients achieved DFR more often when aimed for a DAS<1.6 (18% versus 8%, respectively), suggesting that intensive treatment is helpful in inducing DFR.[35] However, the clinical trials rarely used DFR as a primary outcome and, therefore, the question to what extent the frequency of DFR can be achieved by treatment remains a subject for future studies.

Although we tried to find predictors for DFR, it remains uncertain which patients are able to stop their DMARDs successfully. Meta-analyses could not be performed due to the heterogeneity of studies and effect estimates. Therefore, we summarized and graphically presented data on predictors using predefined criteria, but this

methodology is far less optimal than meta-analysis. Several patients characteristics (e.g. age, SJC, ESR and erosiveness) were not associated with a higher chance of achieving DFR. Results on symptom duration were conflicting, as the relation between DFR and symptom duration was non-significant, but showed a strong tendency towards significance in part of the studies. Furthermore it is known that the association with DFR is not linear but refined to a short period of time[32] (i.e. the window-of-opportunity) and associations may remain undetected if symptom duration is analyzed as a continuous variable. Absence of auto-antibodies was the best predictor for DFR. Although effect sizes were not involved in our analyses, absence of auto-antibodies alone is not sufficient to accurately guide taper-decisions in daily practice. Therefore, effective pursuit of SDFR in clinical practice requires more insight in subsets of patient that are likely to achieve SDFR.

Acknowledging the importance of the auto-antibody status as predictor, the SDFR prevalence will be different for auto-antibody positive and negative patients. We could not stratify the results on SDFR prevalence for autoantibody status as the prevalence reported in the included cohorts and trials was not always stratified for auto-antibodies. However, the studies that included information on autoantibody status in their patient characteristics reported that 52-100% of patients were auto-antibody positive(supplementary table S5).

Since conducting a thorough systematic literature review is time-demanding, a time-gap exist between the actual literature search (March 2019) and publication of the results. As a result, relevant articles in this time interval are not included. A non-systematic screen of articles published in this period revealed that the BioRRA-study[36], published in December 2019. This study focusing on predictors of flare after DMARD-cessation and reported a 52% flare rate (DAS28-CRP≥2.4) after abrupt DMARD-cessation. Predictive of flares were amongst others absence of Boolean remission at baseline, RF-positivity and IL-27. Biomarkers predictive of DFR, as identified in other recent studies, were calprotectin levels and several serum protein levels among which SAA. [37, 38] Calprotectin and SAA are both acute phase reactants. However none of these markers were yet validated in independent studies.

From patients' perspective achieving SDFR is beneficial; it was recently reported to be associated with normalization of functional disability and resolution of symptoms, e.g. fatigue.[28] Unfortunately, clinical trials infrequently evaluated SDFR. If future trials would be designed with DFR/SDFR as primary outcome, consensus of the definition of

remission and the duration of DMARD-free state is required to promote comparability of findings between studies. This may require OMERACT initiatives.

In conclusion, DFR is achievable in RA, and is sustainable in ~10%-20% of patients. DFR can become an important outcome measure for clinical trials, and requires consistency in the definition. Considering the relative short follow-up after DMARD-stop in current clinical trials and the high rate of flares in the first year after DMARD-stop, we propose to incorporate a DMARD-free follow-up of at least 1 year, to ensure that DFR is sustainable.

Supplementary table S1 – Search strategy

(("disease modifying anti rheumatic"[tw] OR "disease modifying anti rheumatoid"[tw] OR "disease modifying antirheumatic"[tw] OR "disease modifying antirheumatoid"[tw] OR "DMARD"[tw] OR "DMARDs"[tw] OR "bDMARD"[tw] OR "bDMARDs"[tw] OR "cDMARD"[tw] OR "cDMARDs"[tw] OR "csDMARD"[tw] OR "csDMARDs"[tw] OR "Antirheumatic Agents"[Mesh:noexp] OR "Antirheumatic Agents"[Pharmacological Action] OR "1-((4,5-bis(4-methoxyphenyl)-2-thiazoyl)carbonyl)-4-methylpiperazine"[tw] OR "1-((4-methylsulfonyl)phenyl)-3-trifluoromethyl-5-(4-fluorophenyl) pyrazole"[tw] OR "1-(4-chlorobenzoyl)-3-(2-(1H-imidazol-1-yl)-2-oxoethyl)-5-methoxy-2-methyl-1H-indole"[tw] OR "2-(4-(quinolin-2-yl-methoxy)phenyl)-2-cyclopentylacetic acid"[tw] OR "2-(4-acetoxyphenyl)-2-chloro-N-methylethylamine"[tw] OR "2-aminomethyl-4-t-butyl-6-iodophenol"[tw] OR "2-diethylaminoethanol"[tw] OR "3-methyl-2-(3-pyridyl)-1-indoleoctanoic acid"[tw] OR "4,5-Dihydro-1-(3-(trifluoromethyl)phenyl)-1H-pyrazol-3-amine"[tw] OR "4-(5-(4-chlorophenyl)-3-(trifluoromethyl)-1H-pyrazol-1-yl)benzenesulfonamide"[tw] OR "4-(acetylamino)benzeneacetic acid"[tw] OR "4-bromo-2,7-dimethoxy-3H-phenothiazin-3-one"[tw] OR "6-(4-fluorophenyl)-2,3-dihydro-5-(4-pyridinyl)imidazo(2,1-b) thiazole"[tw] OR "6-acetylaminocaproic acid"[tw] OR "6-ethoxy-3-(4-methanesulfonylphenyl)-4-phenylpyran-2-one"[tw] OR "7-methoxy-alpha-methyl-2-naphthaleneacetic acid"[tw] OR "A 771726"[tw] OR "Abatacept"[tw] OR "aceclofenac"[tw] OR "acemetacin"[tw] OR "acetaminophen, aspirin, caffeine drug combination"[tw] OR "acetaminophen, butalbital, caffeine drug combination"[tw] OR "acetaminophen, hydrocodone drug combination"[tw] OR "acetosyringone"[tw] OR "acetovanillone"[tw] OR "acetylsalicylic acid lysinate"[tw] OR "Adalimumab"[tw] OR "Adapalene"[tw] OR "Adapalene, Benzoyl Peroxide Drug Combination"[tw] OR "alclofenac"[tw] OR "Allopurinol"[tw] OR "alminoprofen"[tw] OR "alpha-pentyl-3-(2-quinolinylmethoxy)benzenemethanol"[tw] OR "amiprilose"[tw] OR "Ampyrone"[tw] OR "amylase, phosphates, proteases drug combinations"[tw] OR "andrographolide"[tw] OR "anisodamine"[tw] OR "anisodine"[tw] OR "antiflammin P2"[tw] OR "Antipyrine"[tw] OR "Apazone"[tw] OR "apremilast"[tw] OR "Arteparon"[tw] OR "Arthrotec"[tw] OR "Aspirin"[tw] OR "aspirin, aluminum hydroxide, magnesium hydroxide drug combination"[tw] OR "aspirin, butalbital and caffeine drug combination"[tw] OR "aspirin, meprobamate drug combination"[tw] OR "atrinositol"[tw] OR "Auranofin"[tw] OR "Aurothioglucose"[tw] OR "aurotioprol"[tw] OR "Azathioprine"[tw] OR "azulene"[tw] OR "baicalin"[tw] OR "balsalazide"[tw] OR "bendazac"[tw] OR "bendazac lysine"[tw] OR "benorilate"[tw] OR "benoxaprofen"[tw] OR "Benzbromarone"[tw] OR "benziodarone"[tw] OR "benzobarbital"[tw] OR "berbamine"[tw] OR "betulinic acid"[tw] OR "bevonium"[tw] OR "BI 607812 BS"[tw] OR "biphenylylacetic acid"[tw] OR "boldine"[tw] OR "borage oil"[tw] OR "boswellic acid"[tw] OR "bromfenac"[tw] OR "bucillamine"[tw] OR "Bufexamac"[tw] OR "bumadizone"[tw] OR "butibufen"[tw] OR "carbaspirin calcium"[tw] OR "carprofen"[tw] OR "caryophyllene"[tw] OR "castanospermine"[tw] OR "CDP 571"[tw] OR "Celecoxib"[tw] OR "cepharanthine"[tw] OR "Certolizumab Pegol"[tw] OR "Chloroquine"[tw] OR "chloroquine diphosphate"[tw] OR "choline magnesium trisalicylate"[tw] OR "chrysarobin"[tw] OR "Clonixin"[tw] OR "Colchicine"[tw] OR "CP 96345"[tw] OR "Curcumin"[tw] OR "CX 659S"[tw] OR "Cyclophosphamide"[tw] OR "Cyclosporine"[tw] OR "DAB(486)-interleukin 2"[tw] OR "dauricine"[tw] OR "dexketoprofen trometamol"[tw] OR "Diclofenac"[tw] OR "diclofenac hydroxyethylpyrrolidine"[tw] OR "difenpiramide"[tw] OR "Diflunisal"[tw] OR "dimephosphon"[tw] OR "Dipyrone"[tw] OR "diucifon"[tw] OR "droxicam"[tw] OR "DuP 697"[tw] OR "E6011"[tw] OR "ebselen"[tw] OR "ecallantide"[tw] OR "eltenac"[tw] OR "enfenamic acid"[tw] OR "enkephalin-Leu, Ala(2)-Arg(6)-"[tw] OR "Epirizole"[tw] OR "Etanercept"[tw] OR "ethenzamide"[tw] OR "Ethonium"[tw] OR "Etodolac"[tw] OR "etofenamate"[tw] OR "Etoricoxib"[tw]

OR "evening primrose oil"[tw] OR "Febuxostat"[tw] OR "fenamic acid"[tw] OR "fenbufen"[tw] OR "fenclofenac"[tw] OR "fenflumizole"[tw] OR "Fenoprofen"[tw] OR "fentiazac"[tw] OR "fepradinol"[tw] OR "Feprazone"[tw] OR "ferulic acid"[tw] OR "flobufen"[tw] OR "floctafenine"[tw] OR "flosulide"[tw] OR "flunixin"[tw] OR "flunixin meglumine"[tw] OR "flunoxaprofen"[tw] OR "fluproquazone"[tw] OR "Flurbiprofen"[tw] OR "flurbiprofen axetil"[tw] OR "FR 167653"[tw] OR "FR 173657"[tw] OR "Glatiramer Acetate"[tw] OR "glucametacin"[tw] OR "Gold Sodium Thiomalate"[tw] OR "Gold Sodium Thiosulfate"[tw] OR "guacetisal"[tw] OR "Halofenate"[tw] OR "helenalin"[tw] OR "heliodermin"[tw] OR "hemodes"[tw] OR "higenamine"[tw] OR "Hydroxychloroquine"[tw] OR "Ibuprofen"[tw] OR "ibuproxam"[tw] OR "icatibant"[tw] OR "IH 764-3"[tw] OR "imidazole-2-hydroxybenzoate"[tw] OR "indacrinone"[tw] OR "indobufen"[tw] OR "Indomethacin"[tw] OR "Indoprofen"[tw] OR "Infliximab"[tw] OR "Interleukin 1 Receptor Antagonist Protein"[tw] OR "Interleukin-4"[tw] OR "iodoantipyrine"[tw] OR "isoxicam"[tw] OR "kebuzone"[tw] OR "Ketoprofen"[tw] OR "ketoprofen lysine"[tw] OR "Ketorolac"[tw] OR "Ketorolac Tromethamine"[tw] OR "L 745337"[tw] OR "L 778736"[tw] OR "lesinurad"[tw] OR "Levamisole"[tw] OR "licofelone"[tw] OR "lipoxin A4"[tw] OR "lipoxin B4"[tw] OR "lisofylline"[tw] OR "lobenzarit"[tw] OR "lonazolac"[tw] OR "lornoxicam"[tw] OR "loxoprofen"[tw] OR "LQFM-091"[tw] OR "lumiracoxib"[tw] OR "Magnesium Salicylate"[tw] OR "magnolol"[tw] OR "manoalide"[tw] OR "Masoprocol"[tw] OR "mavrilimumab"[tw] OR "Meclofenamic Acid"[tw] OR "Mefenamic Acid"[tw] OR "Meloxicam"[tw] OR "Mesalamine"[tw] OR "Methotrexate"[tw] OR "methyl salicylate"[tw] OR "mizoribine"[tw] OR "MK 473"[tw] OR "mofebutazone"[tw] OR "mofezolac"[tw] OR "N-(2-cyclohexyloxy-4-nitrophenyl) methanesulfonamide"[tw] OR "N-(9H-(2,7-dimethylfluoren-9-ylmethoxy)carbonyl)leucine"[tw] OR "N-succinimidyl-1-(4-chlorobenzoyl)-5-methoxy-2-methyl-1H-indole-3-acetate"[tw] OR "Nabumetone"[tw] OR "nafamostat"[tw] OR "Naproxen"[tw] OR "Nebacetin"[tw] OR "nepafenac"[tw] OR "nifenazone"[tw] OR "Niflumic Acid"[tw] OR "nimesulide"[tw] OR "nitroaspirin"[tw] OR "Olopatadine Hydrochloride"[tw] OR "olsalazine"[tw] OR "olvanil"[tw] OR "oren gedoku to"[tw] OR "orgotein"[tw] OR "oxaceprol"[tw] OR "Oxaprozin"[tw] OR "Oxyphenbutazone"[tw] OR "palmidrol"[tw] OR "parecoxib"[tw] OR "parthenolide"[tw] OR "Penicillamine"[tw] OR "peoniflorin"[tw] OR "phenidone"[tw] OR "Phenylbutazone"[tw] OR "pimecrolimus"[tw] OR "pirfenidone"[tw] OR "Piroxicam"[tw] OR "piroxicam-beta-cyclodextrin"[tw] OR "pirprofen"[tw] OR "Probenecid"[tw] OR "proglumetacin"[tw] OR "propacetamol"[tw] OR "propionylcarnitine"[tw] OR "propyphenazone"[tw] OR "proquazone"[tw] OR "Prospidium"[tw] OR "pyranoprofen"[tw] OR "pyrazolone"[tw] OR "pyrogenal"[tw] OR "rasburicase"[tw] OR "Resveratrol"[tw] OR "rifamycin SV"[tw] OR "Rituximab"[tw] OR "RNS60"[tw] OR "rofecoxib"[tw] OR "rosmarinic acid"[tw] OR "Rumalon"[tw] OR "saiko-keishi-to"[tw] OR "saikosaponin"[tw] OR "salicin"[tw] OR "salicylamide"[tw] OR "Salicylates"[tw] OR "salicylsalicylic acid"[tw] OR "SB 203580"[tw] OR "SC 299"[tw] OR "SC 41930"[tw] OR "SC 560"[tw] OR "semapimod"[tw] OR "seratrodast"[tw] OR "serratiopeptidase"[tw] OR "shikonin"[tw] OR "sinapaldehyde"[tw] OR "sinomenine"[tw] OR "Sodium Salicylate"[tw] OR "ST 679"[tw] OR "Sul-121"[tw] OR "Sulfasalazine"[tw] OR "Sulfinpyrazone"[tw] OR "Sulindac"[tw] OR "sulindac sulfide"[tw] OR "sulindac sulfone"[tw] OR "Suprofen"[tw] OR "suxibuzone"[tw] OR "T0001"[tw] OR "tanshinone"[tw] OR "taxifolin"[tw] OR "tenidap"[tw] OR "tenoxicam"[tw] OR "tepoxalin"[tw] OR "tiaprofenic acid"[tw] OR "tiaramide"[tw] OR "Ticrynafen"[tw] OR "tinoridine"[tw] OR "tisopurine"[tw] OR "tolfenamic acid"[tw] OR "Tolmetin"[tw] OR "tramadol, dexketoprofen drug combination"[tw] OR "tranilast"[tw] OR "traxanox"[tw] OR "tribenoside"[tw] OR "upadacitinib"[tw] OR "ursolic acid"[tw] OR "valdecoxib"[tw] OR "verinurad"[tw] OR "zileuton"[tw] OR "zomepirac"[tw] OR "Zoxazolamine"[tw]) AND ("drug free remission"[tw] OR "drug free remissions"[tw] OR "drug free clinical remission"[tw] OR "drug free disease remission"[tw] OR "DMARD-free remission"[tw] OR (("drug free"[tw] OR "DMARD free"[tw]) AND ("remission"[tw] OR remiss*[tw] OR "Remission Induction"[mesh])) OR "stable remission"[tw] OR "stable remissions"[tw] OR "stable disease remission"[tw] OR "sustained remission"[tw] OR "sustained remissions"[tw] OR "sustained disease remission"[tw] OR "sustained disease remissions"[tw] OR "stable clinical remission"[tw] OR "sustained clinical remission"[tw] OR "sustained clinical remissions"[tw] OR "sustained complete remission"[tw] OR "sustained complete remissions"[tw] OR "stable complete remission"[tw] OR "stable complete remissions"[tw] OR (("stable"[ti] OR "sustained"[ti]) AND ("remission"[ti] OR remiss*[ti] OR "Remission Induction"[mesh])) OR (("drug tapering"[tw] OR "drug taper"[tw] OR "treatment tapering"[tw] OR "medication taper"[tw] OR "medication tapering"[tw] OR "tapering"[tw] OR "taper"[tw] OR "drug discontinuation"[tw] OR "treatment discontinuation"[tw] OR "discontinuation"[tw] OR "cessation"[tw]) AND ("remission"[tw] OR remiss*[tw] OR "Remission Induction"[mesh]))) AND ("Arthritis, Rheumatoid"[Mesh:noexp] OR "rheumatoid arthritis"[tw] OR (rheumatoid*[tw] AND arthriti*[tw]))) NOT (("Case Reports"[ptyp] OR "case report"[ti] OR "Editorial"[ptyp] OR "Comment"[ptyp]) NOT ("Review"[ptyp] OR "Clinical Study"[ptyp])) AND (english[la] OR dutch[la])

Supplementary table S2 – In- and exclusion criteria study selection

Inclusion criteria (1) Study designs which will be included are all observational studies, including cross-sectional, case-control, prospective and retrospective cohort studies, and intervention studies, controlled and uncontrolled. (2) Patients with rheumatoid arthritis will be included, in which patients should meet the ACR/EULAR 1987 and/or 2010 criteria for RA or RA should been diagnosed by a Rheumatologist (expert opinion). (3) The study should clearly report on DMARD-free remission, i.e. complete discontinuation of all DMARDs, including glucocorticoids is required. The study should report clearly on tapering and complete discontinuation of all DMARDs until eventually a complete DMARD-free state will be achieved. (5) Only studies in English, and Dutch, with available full text will be included. (6) Only studies with full paper articles available will be included. (7) All years of publication will be included.
Exclusion criteria (1) Case-report studies and reviews will be excluded. Reviews were screened for (extra) eligible studies. (2) Studies focussing on disease other than rheumatoid arthritis (also unspecified arthritis) will be excluded. However, studies focussing on other disease next to RA can be included from which the RA data will only be extracted (3) Studies in which the use of not all DMARDs are tapered and completely discontinued. Studies which do not clearly describe if all DMARDs are completely stopped, i.e. no transparency of the DMARD-free state, will also be excluded. (4) Studies should focussing on other reasons for tapering and discontinuation of DMARD then remission, e.g. adverse events, retention etc. will be excluded.

Legend: In- and exclusion criteria used for study selection. ACR: American College of Rheumatology, EULAR: European League Against Rheumatism, RA: Rheumatoid arthritis, DMARD: Disease -modifying antirheumatic drug.

Supplementary table S3 – Data extraction form

(Name of study)
Articles included
1. ....
2. ....
etc.
DATA- EXTRACTION form Acronym study
Study design: clinical trial/observational

Supplementary table S3(Continued)

(Name of study)
Country
Years of inclusion
Inclusion criteria
Exclusion criteria
RA criteria used
Study population (n)
Baseline characteristics (patient characteristics)
Primary outcome
Secondary outcomes
Intervention (arms) (if applicable)
Medication prescribed during study
Glucocorticosteroids use (y/n, dosage)
Follow-up (FU)
Monitoring during FU
Lost tot FU
Remission/DFR specific data extraction
Remission criteria
Tapering (start/cut-off point)
Tapering methods
DFR definition specified (y/n)
DFR criteria used
DFR duration reported
Sustained DFR reported (y/n)
Flare definition
Quantitative data-extraction
Remission (n)
Tapering (n)
DFR (n)
SDFR (n)
Flares (n)
Remarks (in general)

Legend: Data extraction form used for systematic data-extraction Data extraction forms were fulfilled independently by two reviewers (MV and EvM), disagreements were discussed until consensus was reached. DFR: DMARD-free remission, FU: Follow-up.

Supplementary table S4 – Risk assessment tool

QUALITY ASSESSMENT	
DMARD-free remission	(+) (-) (?)
1. DFR definition *(description of DFR criteria/definition)* 2. DFR duration *(description of period between DMARD-stop and being appoint as DFR)*	
General study quality	(+) (-) (?)
Study population	
1. Selection of patients *(description of in-/exclusion criteria)* 2. Criteria used for RA diagnosis	
3. Baseline characteristics study population *(description of characteristics)*	
Randomization	
4. Randomization for different study treatments	
Blinding (combined score)	
5.1. Blinding outcome assessors	
5.2. Blinding patients	
Interventions	
6. Treatment strategies *(description of strategies)* 7. Cut-off point tapering *(description of cut-off point)* 8. Tapering methods *(description of methods)*	
Follow-Up	
9. Organisation of follow-up *(frequency of monitoring)* 10. Lost-to-follow-up	
Analysis & Data presentation	
11. Outcome reporting	
12. Analysis techniques *(description of techniques)* 13. Missing data *(handling of missing data described)*	

Legend: DMARD: Disease-modifying antirheumatic drug, DFR: DMARD-free remission, RA: Rheumatoid arthritis.

Supplementary table S5 - Summary of studies reporting on discontinuation of DMARDs in rheumatoid arthritis

Clinical trials

Study	Inclusion period	Study pop (n)	Baseline characteristics	Intervention	Follow Up (years)	Monitoring during FU	LTFU	Remission criteria	Tapering (initiation)	Tapering Method (duration of tapering)	GCs use
ACT-RAY	2009 - 2013	556	♀ 80.3%, age 53.3y RA 1987 criteria, RA duration 8.3y DAS28 6.35 at BL Unresponsive to MTx	0-52w: TCZ (277) or TCZ+MTx (279) 52-104w: DAS28<2.6 = Tapering (472) DAS28 2.6 – 3.2 = Continue DAS28 >3.2 = + csDMARDs	2	Every 12w	138	DAS28<2.6	DAS<2.6 at 2 visits	1st bDMARDs 2nd c sDMARDs (12-36w)	GCs tapered to 5mg/d after 24w. Tapering in drug-free state not reported.
AVERT	2010 - 2014	351	♀ 77.8%, age 47y (median) No spec. RA criteria All patients ACPA+, 95% RF+ Symp duration 0.56y DAS28-CRP 5.4 at BL	0-52w: ABA+MTx or ABA or MTx 52-72w: Tapering if DAS28<3.2	1.5	Every 3m	12	DAS28-CRP<2.6	DAS<3.2 After 52 w ABA+MTx or ABA or MTx	bDMARD abrupt, MTx + GCs gradual (1 month)	GCS tapered after 12m in 1 month time.

Supplementary table S5 (Continued)

Clinical trials

Study	Inclusion period	Study pop (n)	Baseline characteristics	Intervention	Follow Up (years)	Monitoring during FU	LTFU	Remission criteria	Tapering (initiation)	Tapering Method (duration of tapering)	GCs use
BeSt	2000 - 2002	508	♀ 68%, age 54y 1987 RA criteria RF+ 65% Symp duration 23-26w (median) DAS44 4.3-4.5 at BL	Arm1 = Monotherapy Arm2 = Step-up therapy Arm3 = csDMARDs combi-therapy Arm4 = MTx+TOCI	5	Every 3m	195 (38%)	DAS44<1.6	DAS44<1.6 min 6m	One by one & gradual tapering (time for taper n.r.)	i.a. allowed
IMPROVED	2007 - 2010	479	♀ 68%, age 52y 2010 RA criteria ACPA 55%, RF+ 56% Symp duration 18w (median) DAS44 3.2±0.91**	0-4m: MTx (+GCs bridging) >4m: - DAS44<1.6 = Tapering - DAS44>1.6: Arm1 = MTx + bDMARDs Arm2 = MTx +csDMARDs	5	Every 4m	112	DAS44<1.6 or Boolean remission (ACR 2011)	DAS44<1.6	Gradual (MTx in 10w, tapering times other DMARDs n.r.)	GCs tapered.
PRIZE	2009 - 2012	306	♀ 64.8%, age 49y 1987 RA criteria ACPA+ 65.8% Symp duration 6.8	0-52w: 50 mg ETA+MTx 52-91w: 25 mg ETA+MTx or MTx or PBO 91-117w: Tapering if DAS28<3.2	±2 (27m)	Every 3m	97	DAS28<2.6	DAS28<3.2	52-91w: 50% bDMARD >91w: Abrupt stop bDMARDS, than taper	GCs tapered, complete withdrawal at 39w

Supplementary table S5 (Continued)

Clinical trials

Study	Inclusion period	Study pop (n)	Baseline characteristics	Intervention	Follow Up (years)	Monitoring during FU	LTFU	Remission criteria	Tapering (initiation)	Tapering Method (duration of tapering)	GCs use
			DAS28 5.8 at BL							csDMARDs in 2-4w (total tapering time 41-43w)	
RETRO	2010 - 2013	101	♀ 62%, age 57y (median) 2010 RA criteria ACPA+ 60% RA duration 5y DAS28 1.91 at BL	Arm 1 = Continuation DMARDs Arm 2 = Tapering DMARDs Arm 3 = Discontinuation DMARDs	1	Every 3m	n.r.	DAS28<2.6	DAS28< 2.6 for min 6m (3 visits)	Gradual: 6m 50% dose reduction than stop. (6m tapering)	GCs tapered
tREACH	2007 - 2011	281	♀ 68%, age 53.2y 2010 RA criteria (95%) ACPA 80% Symp duration 166 days DAS44 3.36 at BL (mean)	Arm 1 = Triple therapy (MTx+SSZ+HCQ) Arm 2 = MTx monotherapy	2	Every 3m	76	DAS44<1.6 at 2 visits	DAS44<1.6 at 2 visits	Gradual: Dose reduction 50% → 25% → stop (6m tapering)	Bridging at start study, afterwards GCs tapered.

Supplementary table S5 (Continued)

Clinical trials

Study	Inclusion period	Study pop (n)	Baseline characteristics	Intervention	Follow Up (years)	Monitoring during FU	LTFU	Remission criteria	Tapering (initiation)	Tapering Method (duration of tapering)	GCs use
U-Act-Early	2010 - 2012	317	♀67%, age 54y 1987/2010 RA criteria ACPA 70% DAS28 5.2 at BL	Arm 1 = TOCI+MTx (106) Arm 2= TOCI (103) Arm 3 = MTx (108)	2	No fixed visits, ± every 4w	80	DAS28<2.6 & SJC ≤4	DAS28<2.6 & SJC ≤4 for 24w (= sustained remission)	MTx gradual 5mg/w until 10 mg →stop (1m no tapering) bDMARDs to 50% 3m àstop (total tapering time ±4-6m)	Incidental GCS i.a. allowed, and p.o. once a year max 2w
Brocq et al	1999 - 2005	21	♀61.9%, age 61y 1987 RA criteria ACPA 53% RA duration 11.3y Rem duration 19.2m	Discontinuation of TNFi	1	Every 1m	1	DAS28<2.6 min 6m	DAS<2.6 min 6m (inclusion criterion)	Abrupt DMARD stop	3 patients used GCs, max 5mg/d

Supplementary table S5 (Continued)

Clinical trials

Study	Inclusion period	Study pop (n)	Baseline characteristics	Intervention	Follow Up (years)	Monitoring during FU	LTFU	Remission criteria	Tapering (initiation)	Tapering Method (duration of tapering)	GCs use
El Miedany et al.	Not reported <2016	157	♀ 61.4%, age n.r. 2010 RA criteria Use of stable cs+bDMARDs ACPA+ 62% DAS28 2.08 at BL	Arm 1 = bDMARDs 50% (32) Arm 2 = cs+bDMARDs 50% (32) Arm 3 = bDMARDs stop (32) Arm 4 = cs+bDMARDs stop (32) Arm 5 = continuation (32)	1	Every 1m	3	DAS28<2.6 min 6m	Protocol (arm 4)	Abrupt DMARD stop (arm 3+4)	GCs were not allowed in study.
SURPRISE	2009 - 2012	233	♀ 89%, age 55.9y 1987 RA criteria RF+ 72.5% Disease duration 3.5y DAS28 5.0	Arm 1 = TOCI + MTx (ADD-ON) Arm 2 = TOCI (SWITCH)	2	Every 3m	4	DAS28<2.6	DAS28<2.6	Abrupt DMARD stop	GCs allowed, 15.1% used GCs (mean 3.4mg/d)
Kita et al.	2008 - 2009	13	♀ 69.2%, age 59.2y 2010 RA criteria RA duration 13.7w SDAI 20.2 at BL	0-12m: Treat-to-target > 12m: Tapering if SDAI remission	1	Every 3m	1	SDAI	SDAI for 12m + 33% reduction BME on MRI	Assumable abrupt DMARD stop	GCs allowed at BL, unclear whether GCs were tapered.

Supplementary table S5 (Continued)

Clinical trials

Study	Inclusion period	Study pop (n)	Baseline characteristics	Intervention	Follow Up (years)	Monitoring during FU	LTFU	Remission criteria	Tapering (initiation)	Tapering Method (duration of tapering)	GCs use
Ten Wolde et al.	n.r. <1997	285	ACPA+/RF+ 100% ♀ 58%, age 61.1y 1987 RA criteria RA duration 10.6y ARA remission (5/6 criteria) at BL RF+ 70%	Arm 1 = Continue treatment Arm 2 = Switch to placebo (abrupt)	1	4, 8, 12, 26 52w	9	ARA remissionA (5 out of 6)	Protocol	Abrupt switch to placebo	GCs not allowed.
DREAM trial	2008 - 2010	187	♀ 87.7%, age 57y (median) 1987 RA criteria RA duration 7.8y (median) DAS28 1.5 (median)	Discontinuation TCZ	1	Every 1m	161	DAS28<2.6	DAS28<3.2	Abrupt DMARD stop	GCs allowed, at start of tapering (dose 0-7.0 mg/d)

Supplementary table S5 (Continued)

Observational studies

Study	Inclusion period	Study pop (n)	Baseline characteristics	Intervention	Follow Up (years)	Monitoring during FU	LTFU	Remission criteria	Tapering (initiation)	Tapering Method (duration of tapering)	GCs use
DREAM cohort	2006 - 2009	229	♀ 63.3%, age 57.5y RA expert opinion (79% 1987 RA) Symp duration (median) 13w ACPA 58.6 DAS28 4.9	Treat-to-target, steered at DAS28<2.6: initial MTx monotherapy, if DAS28>2.6 + SSZ if DAS28>3.2 TNF-inhibitor	5	Every 3m	58	DAS28<2.6	DAS28<2.6 for 6m	Gradual (duration of tapering n.r.)	<10mg/d, i.a. permitted
Tiippanna - Kinnunen et al	1986 - 1989	70	♀ 79.3%, age 44y 1987 RA criteria RA duration 8m RF 65%	Saw-tooth strategy: gold, SSZ of HCQ	15	Year 1-3: every 3-4 months, then visits at: 5y, 7y, 10y and 15y	17***	ARA remission (5/6 criteria)	'Clinical remissionB min 12m' or 'symptom-free & prolonged minor disease activity'	n.r.	<10mg in active disease, in remission i.a. allowed

Supplementary table S5 (Continued)

Observational studies

Study	Inclusion period	Study pop (n)	Baseline characteristics	Intervention	Follow Up (years)	Monitoring during FU	LTFU	Remission criteria	Tapering (initiation)	Tapering Method (duration of tapering)	GCs use
Leiden EAC	1993 - 2011	889	♀ 66.7%, age 56.5y 1987 RA criteria Symp duration (median) 4.4 m ACPA 52.4% SJC 6-12 (median)	'93-'95 NSAIDs '96-'98 Mild DMARDs (HCQ, SSZ) '99-'02 Initial treatment: MTx/SSZ > 2002 Treat-to-target: initial MTx	1-19	BL, 4 months, annually	n.r.	Remission criteria not specified. DFR: No clinical synovitis	n.r.	Not protocolized	Not allowed by definition of DFR
ESPOIR	2002 - 2005	533	♀ 75.8%, age 48.8y RA expert opinion Symp duration (median) 21.3w SJC 7 (median)	Treated with csDMARDs	5	Every 6 months first 2 years, afterwards annually	n.r.	n.r.	n.r.	n.r.	Unknown
ERAS	1986 - 1996	895	♀ 69% age 52y ACR criteria Symp duration 8.3m RF 63% SJC 14 (median)	Rheumatologist preference, predominantly MTx, SSZ, HCQ	10	Every 4m, at some point annually	384*** (30%)	n.r.	n.r.	n.r.	Unknown

LTFU: Loss to follow-up

Legend:

Data is presented as mean, unless otherwise indicated.

** Both the UA and RA population together (n=610), *** Due to retrospective design, LTFU not taken into account as part of the study population.

[A] ARA remission: morning stiffness absent (or not exceeding 15 minutes), no fatigue, no joint pain by history, no joint tenderness, no joint or tendon sheath swelling, no elevation of ESR (in 5/6, fatigue is not included in the criteria). [B] Clinical remission defined as no tender joints , no swollen joints, no joint pain by history, ESR<30(female/<20(male) for minimal 12 months. Or prolonged symptom-free phase of disease with minor disease activity.

ACR: American College of Rheumatology, ADA: Adalimumab, BL: Baseline, bDMARDs: biological DMARDs, crit: criteria, csDMARDs: conventional DMARDs, DAS: Disease activity score, DFR: DMARD-free remission, Establ.: Established, ETA: Etanercept, FU: Follow-Up, GCs: glucocorticosteroids, HCQ: Hydroxychloroquine, HDA: High disease activity, IFX: Infliximab, n.r.: not reported, LDA: Low disease activity, rem: remission, MTx: Methotrexate, m: months, RA: rheumatoid arthritis, symp: symptom, SSZ: Sulfasalazine, n.r.: not reported, TCZ: Tocilizumab, TNFi: TNF-α inhibitor, w: weeks, y: years.

Supplementary table S6 – All predictors of DFR retrieved from the literature

CLINICAL BIOMARKERS

×	**Age**	✓
v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 1.01(0.98-1.03)*	no associations between age and DFR have been reported within the included articles	
v.d. Woude et al. (2009) (LEAC, n=454)[C] *HR 1.02(0.99-1.03)*		
Kuijper et al. (2016)(tREACH, n=281)[L] *OR 0.995(CI not specified)*		
v.d. Woude et al. (2012) (LEAC, n=424) [L] *OR 1.02(0.998-1.04)*		
v.d. Woude et al. (2009) (ERAS, n=895) [C] *HR 1.00(0.98-1.01)*		
de Rooy et al. (2011) (LEAC, n=676) [L] *OR 0.99(0.97-1.00)* *(not achieving DFR)*		
v.d. Kooij et al. (2009) (BeSt, n=508)[B] *nDFR 54y, DFR 56y*		
Ajeganova et al. (2016)(LEAC, n=886) *nDFR 57.4y, DFR 56.3y* [B]		
Emery et al. (2018)(PRIZE, n=65)[L] *(no estimates specified)*		

×	**Gender**	✓	
v.d. Woude et al. (2009) (LEAC, n=454)[C] *Female: HR 1.28(0.74-2.19)*	Kuijper et al. (2016)(tREACH, n=281)[L] *Female: OR 0.352* [M] (CI not specified)*		↑
v.d. Woude et al. (2009) (ERAS, n=895)[C] *Female: HR 0.78(0.50-1.2)*	v.d. Woude et al. (2012) (BeSt, n=508) [L] *Male: OR 2.39(1.26-4.53)* [M]*		↑
de Rooy et al. (2011)(LEAC) [L] *Female, OR 0.85 90.50-1.45)* *(not achieving DFR)*	v.d. Kooij et al. (2008)(BeSt) [L] *Male:DFR 52% vs nDFR 29%*[M]* *(OR not specified)*		↑
v.d. Woude et al. (2012) (LEAC, n=424)L *Female, OR 1.19(0.62-2.28)*			
Nishimoto et al. (2014) (DREAM, n=187)[C]			

Supplementary table S6 (Continued)

CLINICAL BIOMARKERS

Male HR 0.61(0.37-1.00)			
Ajeganova et al.(2016)(LEAC, n=886)[B] *Female: nDFR63%, DFR 68%*			
Emery et al. (2018)(PRIZE, n=65)[L] *Male (no estimates specified)*			
×	**BMI**	✓	
v.d. Woude et al. (2012) (LEAC, n=424)[L] *OR 0.95(0.83-1.08)*	de Rooy et al. (2011)(LEAC) [L] *OR 1.11(1.01-1.23)*[,U] (not achieving DFR)*		↓
v.d. Woude et al. (2012) (BeSt, n=508)[L] *O R0.96(0.88-1.04)*			
v.d. Woude et al. (2009) (ERAS, n=895)[C] *HR 0.98(0.93-1.04)*			
×	**Smoking**	✓	
v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 0.69(0.36-1.33)*	v.d. Woude et al. (2009) (LEAC, n=454)[C] *HR 0.56(0.34-0.94)* U*		↓
v.d. Woude et al. (2009) (ERAS, n=895)[C] *HR 0.54(0.29-1.02)*	v.d. Woude et al (2012) (LEAC, n=424) [L] *OR 0.48(0.25-0.93)*[U]*		↓
Ajeganova et al. (2016)(LEAC, n=886) Smoking ever: nDFR 54%, DFR 56% [B]			
×	**Family history of RA**	✓	
v.d. Woude et al. (2009) (LEAC, n=454)[C] *HR 0.55(0.30-1.04)*	de Rooy et al. (2011) (LEAC, n=676) [L] *OR 2.27(1.18-4.36)* [U] (not achieving DFR)*		↓
v.d. Woude et al. (2009) (ERAS, n=895)[C] *HR 0.87(0.53-1.44)*			
×	**Miscellaneous**	✓	
v.d. Woude et al. (2009) (LEAC, n=454) [C] *Absence comorbidities* *HR 0.98(0.59-1.61)* Kuijper et al. (2016)(tREACH, n=281) [L] *Dutch ethnicity OR 3.316*[U] *(CI not specified)*	Kuijper et al. (2016)(tREACH, n=281) [L] *Paid work: OR 0.438* [M] (CI not specified)*		↓

Supplementary table S5 (Continued)

CLINICAL BIOMARKERS

×	(shorter) Symptom duration	✓	
v.d. Woude et al. (2012)(BeSt, n=508) (cont) [C] *O R0.99(0.98-1.00) (cont. in weeks)*	v.d. Linden et al. (2010) (LEAC, n=598) [C] ≥12w symptoms vs <12w *HR 1.90* [M] (1.18 - 3.05)(not achieving DFR)*		↑
Kuijper et al. (2016)(tREACH, n=281) [L] *OR 1.00 (CI not specified) (measure symp duration not specified)*	v.d. Woude et al. (2009) (ERAS, n=895) [C] *HR 0.94*M(0.89-0.99)* *(continuous in months)*		↑
Akdemir et al. (2018) (BeSt/IMPROVED, n=133/175)[L] *OR 0.98 (0.97-1.00)* *(continuous in weeks)*	v.d. Woude et al. (2009) (LEAC, n=454) [C] *HR 0.94(0.88-0.99)* [U]* *(continuous in months)*		↑
Emery et al. (2018)(PRIZE, n=65)[L] *(no estimates specified)*	v.d. Kooij et al. (2009)(BeSt, n=508) [L] *DFR:18w(11-33), nDFR:24w(14-56)**[M] (continuous) *(OR not specified)*		↑
Nishimoto et al.(2014) (DREAM, n=187)[C] *Disease duration (<7.8y vs >7.8y (median)) HR 0.81(0.60-1.00)*	de Rooy et al. (2011)(LEAC, n=676) [L] *OR 1.02(1.01-1.03)* [L, U]* *(continuous in weeks)* *(not achieving DFR)*		↑
	v.d. Woude et al. (2012) (LEAC, n=424) [L] *OR0.98(0.96-0.99)* [U] (cont. in weeks)*		↑
	Ajeganova et al. (2016)(LEAC, n=886) [B] *nDFR:4.7(2.4-8.6), DFR:2.9* *(1.8-6.5)** (continuous in months)*		↑
×	**Disease activity score at baseline**	✓	
Akdemir et al. (2018) (BeSt/IMPROVED, n=133/175)[L] *OR 0.94(0.58-1.53)* *Only IMPROVED data selected for figure*	v.d. Woude et al. (2012)(BeSt. N=508) [L] *OR 0.63(0.43-0.94)* [M]*		↓
Emery et al. (2018)(PRIZE, n=65)[L] *(no estimates specified)*	Kuijper et al. (2016)(tREACH, n=281) [L] *OR 0.587* [M] (CI not specified)*		↓
	v.d. Woude et al. (2009) (ERAS, n=895) [C] *HR 0.65**(0.55-0.76) [U]*		↓
	v.d. Kooij et al. (2008)(BeSt, n=508) [B] *nDFR 4.5 vs. DFR 4.1**		↓
	Nishimoto et al. (2014)(DREAM, n=187) [C] *HR 0.59 (0.44-0.81) [U]*		↓
×	**Swollen Joint Count at baseline**	✓	
v.d. Woude et al. (2009) (LEAC, n=454)C – 44-SJC *HR 1.00(0.96-1.04)*	v.d. Woude et al. (2009) (ERAS, n=895) – 44-SJC *HR 0.97*(0.95-0.99) C,M*		↓

Supplementary table S6 (Continued)

CLINICAL BIOMARKERS			
v.d. Woude et al. (2012)(BeSt, n=508)[L] *OR 1.01(0.97-1.06)*			
v.d. Woude et al. (2012) (LEAC, n=424)[L] *OR 0.99(0.94-1.04)*			
de Rooy et al. (2011)(LEAC, n=676)[L] *OR 0.99(0.96-1.02)(not achieving DFR)*			
Ajeganova et al. (2016) (LEAC, n=886)[B] – 66-SJC *nDFR 8(4-15), DFR 8(4-13)*			
Emery et al. (2018)(PRIZE, n=65)[L] *(no estimates specified)*			
×	**Tender Joint Count at baseline**	✓	
Ajeganova et al. (2016) (LEAC, n=886)[B] – 68-TJC *nDFR 7(5-11), DFR 8(5-11)*		v.d. Woude et al. (2009) (ERAS, n=895) - RAI [C] *HR 0.92**(0.88-0.97)* [M]	↓
×	**Morning stiffness**	✓	
v. Nies et al. (2015)(LEAC, n=807)[C] *HR 0.85(0.65-1.11)*		v.d. Kooij et al. (2008)(BeSt, n=508)[B] *VAS morning stiffness: nDFR 60(24) vs DFR 54(24)**	↓
v. Nies et al. (2015)(ESPOIR, n=353)[C] *HR 0.80(0.50-1.29)*			
de Rooy et al (2011)(LEAC, n=676)[L] *OR 1.00(0.99-1.01) (not achieving DFR)*			
Ajeganova et al. (2016)(LEAC, n=886)[B] *Morning stiffness VAS: nDFR 64(36-81), DFR 57(36-76)*			
×	**Miscellaneous**	✓	
v.d. Woude et al. (2009) (LEAC, n=454)[C] *Acute onset HR1.55(0.94-2.56)* *Onset in small joints HR1.48(0.91-2.40)* *Onset symmetrical symptoms HR1.24(0.72-2.14)*		v.d. Woude et al. (2009) *(ERAS, n=895) - Acute onset symp* [C] *HR 2.03*(1.15-3.59)* [M]	
v.d. Woude et al. (2009) (ERAS, n=895) [C] *Start small joints HR1.27(0.80-2.04)* *Start symm sympt 1.18(0.67-2.07)*		Burgers et al. (2018)(LEAC) [C] *LJI HR1.4(1.0-2.0)** [M]	↑
de Rooy et al. (2011) (LEAC, n=676)[L] *Chronic vs acute OR1.55(0.93-2.59)*			↑
Small vs Large joints *OR 0.66(0.34-1.28)* *Upper vs lower extremitites*			

Supplementary table S6 (Continued)

CLINICAL BIOMARKERS		
OR 0.76(0.35-1.62) *Upper and lower vs lower extremities OR 1.01(0.47-2.26)* *Symm vs asymm symptoms OR 0.89(0.51-1.55)*		
×	**HAQ**	✓
v.d. Woude et al. (2009) (LEAC, n=454)[C] - m-HAQ *HR 1.06(0.74-1.52)*	v.d. Woude et al. (2012) (BeSt, n=508) [L] *OR 0.63(0.40-0.98)** [U]	↓
v.d. Woude et al. (2012) (LEAC, n=424)[B] *OR 1.26(0.78-2.03)*	Kuijper et al. (2016)(tREACH, n=281) [L] *OR 0.515** [U]	
Ajeganova et al. (2016) (LEAC, n=886)[B] *nDFR 1.0(0.63-1.50), DFR 1.0(0.62-1.50)*	v.d. Woude et al. (2009)(ERAS, n=895) - m-HAQ [C] *HR 0.66*(0.44-0.99)* [M]	↓
Emery et al. (2018)(PRIZE, n=65)[L] *(no estimates specified)*	v.d. Kooij et al. (2008)(BeSt, n=508) [B] *nDFR 1.4 vs. DFR 1.2**	↓
	Nishimoto et al. (2014)(DREAM, n=187) [C] *HR 0.73(0.53-0.99)*[U]	↓
×	**Visual Analogue Scale**	✓
Ajeganova et al. (2016) - VAS pain [B] *nDFR 52(34-70), DFR 48(29-65)*	v.d. Kooij et al. (2008)(BeSt, n=508) *VAS pain: nDFR 55 vs DFR 45*** [B] *VAS disease activity: nDFR 61(23) vs DFR 55(19)** [B]	↓
Emery et al. (2018)(PRIZE, n=65) -VAS pain [B] *(no estimates specified)*	Ajeganova et al. (2016)(LEAC, n=886) [B] *VAS patient: nDFR55(34-76),* *DFR:51(33-67)** *VAS fatigue: nDFR:50(17-70),* *DFR:40(12-60)**	↓
×	**Miscellaneous**	✓
Kuijper et al. (2016) (tREACH, n=281) – SF36 [L] *OR 1.056(CI not specified)* Emery et al. (2018) (PRIZE) - SF36/mTSS/SGA score [L] *(no estimates specified)* Nishimoto et al. (2014) (DREAM, n=187) *Steinbrocker stageC HR 0.77(0.57-1.04)* *Steinbrocker class C HR 0.82(0.20-3.33)* Kuijper et al. (2016)(tREACH, n=281) [L] *OR 1.399 (CI not specified)*	v.d. Linden et al. (2009) (LEAC, n=687) [C] *HR 4.7 (2.8-8.0)** [U] *(for not achieving DFR)*	

Supplementary table S6 (Continued)

LABORATORY BIOMARKERS

×	**Reuma factor**	✓
		v.d. Woude et al. (2012)(BeSt, n=508) [L] *OR 0.39(0.21-0.70)** [U]
		v.d. Woude et al. (2009) (ERAS, n=895) [C] *HR0.28**(0.16-0.49),* [M]
		v.d. Woude et al. (2009) (LEAC, n=454) [C] *HR 0.17(0.10-0.31)* [U]**
		de Rooy et al. (2011) (LEAC, n=676) [L] *OR6.66(3.69-12.02)** [U] *(not achieving DFR)*
		v.d. Linden et al. (2011)(LEAC) [C] *RF level >3x ref: HR 5.7(2.9-11.4)** [C, M] *RF50 HR 3.1(1.2-7.6)* [M]
		v.d. Woude et al. (2012) (LEAC, n=424) [L] *OR 0.22(0.11-0.44)** [U]
		Ajeganova et al. (2016)(LEAC, n=886) [B] *nDFR:65% RF+, DFR:31% RF+*** *(low or high ACPA positive not isgn different nDFR/DFR)*
		v.d. Kooij et al. (2008)(BeSt, n=508) [B] *IgM RF neg: nDFR 33% vs DFR 48%**
		Nishimoto et al. (2014) (DREAM, n=187) [C] *HR 0.53(0.33-0.85)* [U]
		Emery et al. (2018)(PRIZE, n=65)[L,U]* *(no estimates specified)*
×	**Anti-CCP**	✓
Kuijper et al. (2016)(tREACH, n=281) [L] *OR 0.636 (CI not specified)*		v.d. Woude et al. (2012)(BeSt, n=508) [L] *OR 0.20(0.10-0.39)** [M] v.d. Linden et al. (2009) (LEAC), n=687) [C]
		v.d. Linden et al. (2009) (LEAC), n=687) [C] *anti-CCP-2 HR 11.6(5.8-23.4)* [U] *anti-CCP3 HR 6.0(3.4-10.4)* [U] *(for not achieving DFR)*
		v.d. Woude et al. (2009) (LEAC, n=454) [C] *HR0.09(0.04-0.20)** [M]

Supplementary table S6 (Continued)

	LABORATORY BIOMARKERS	
		v.d. Linden et al. (2011) (LEAC, n=598) [C] *HR 11.3 (5.6-22.7)*** [M] *(not achieving DFR)*
		de Rooy et al. (2011)(LEAC, n=676) [L] *OR11.46(5.85-22.46)*** [U] *(not achiev*ing *DFR)*
		v.d. Kooij et al. (2008)(BeSt, n=508) [L] *Anti-CCP neg: DFR 57% vs nDFR 36%** [M] *(OR not specified)*
		Ajeganova et al. (2016) (LEAC, n=886) [B] nDFR 62% ACPA+, DFR 18% ACPA+**
		Emery et al. (2018)(PRIZE, n=65)[L,U]* *(no estimates specified)*
		v.d. Broek et al.(2012)(BeSt, n=484) [L] *RR 0.4(0.3-0.7)*,* [M]
×	**Anti-MCV**	✓
		v.d. Linden et al. (2009) [C] *HR 4.9 (3.0-8.2)* [U] *(not achieving DFR)*
		de Rooy et al. (2011, n=676) [L] *OR 6.13(3.48-10.79)** [U] *(not achieving DFR)*
×	**CRP**	✓
v.d. Woude et al. (2012) (BeSt, n=508) [L] *OR 1.00(0.99-1.01)*		v.d. Woude et al. (2009) (LEAC, n=454) [C] *HR 0.99(0.98-1.0)** [M]
v.d. Woude et al. (2012)(LEAC, n=424)[L] *OR 1.00(0.99-1.01)*		
de Rooy et al. (2011)(LEAC, n=676)[L] *OR 1.01 (0.997-1.1013) (not achieving DFR)* v. Steenbergen et al. (2015)(LEAC, n=645) *rs1896368 (DKK-1)/rs1896367/rs1528873*		
Ajeganova et al. (2016)(LEAC, n=886)[B] *nDFR 15(6-38), DFR 16(16-33)*		
Emery et al. (2018)(PRIZE, n=65)[L] *(no estimates specified)*		
×	**ESR**	✓
v.d. Woude et al. (2009) (LEAC, n=454)[C] *HR 0.99(0.98-1.00)*		

Supplementary table S6 (Continued)

	LABORATORY BIOMARKERS	
v.d. Woude et al. (2012)(BeSt, n=508) *OR 0.99(-0.98-1.00)* v.d. Woude et al. (2009) (ERAS, n=895)[C] *HR 0.99(0.99-1.00)* de Rooy et al. (2011)(LEAC, n=676)[L] *OR 1.01(0.995-1.015)* v.d. Woude et al. (2012)(LEAC, n=424)[L] *OR 1.00(0.99-1.01)* Ajeganova et al. (2016)(LEAC, n=886)[B] *nDFR 32(18-53), DFR 29(16-48)* Emery et al. (2018)(PRIZE, n=65)[L] *(no estimates specified)*		
×	**IL-2**	✓
		v. Steenbergen et al. (2015) (LEAC) - serum IL2Rx levels [C] *Lower IL2 levels: HR0.83(0.70-0.98)** [U]
×	**IL-6**	✓
		Nishimoto et al. (2014) (DREAM, n=187) [C] *IL-6 (<35pg/ml vs >35pg/ml) HR 0.41 (0.27-0.63) **[M]*
×	**MMP-3**	✓
		Nishimoto et al. (2014)(DREAM, n=187) C MMP-3 (normal vs abnormal) HR0.29(0.19-0.43) *M
×	**Shared Epitope**	✓
v. Heemst et al. (2015)(LEAC, n=441) *HLA DRB1*13 higher chance DFR*, but after stratification for ACPA status was this effect no longer present.*		v.d. Woude et al. (2009) (LEAC, n=454) - HLA [C] *HR 0.46 (0.29-.75)*,[U]*
/rs26232(C5orf30)/rs11908352(MMP-9)/ rs451066/rs1485305 (OPG)		v.d. Woude et al. (2012)(BeSt, n=508) [L] *OR 0.46(0.25-0.85)*[U]*
		v.d. Woude et al. (2009) (ERAS, n=895) - HLA [C] *HR 0.44(0.26-0.73)*[M]*
		de Rooy et al. (2011) (LEAC, n=676)- HLA [L] *OR2.25(1.35-3.74)**,[U] *(not achieving DFR)*

Supplementary table S6 (Continued)

LABORATORY BIOMARKERS

		v.d. Woude et al. (2012) (LEAC, n=424) [L] *OR0.35(0.19-0.66)*[U]*
×	**Other**	✓
de Rooy et al. (2011)(LEAC, n=676)[L] *CD40 non-G carrier OR 0.78(0.17-3.54)*		v.d. Linden et al. (2009) *Combinations of auto-antibodies* *anti-CCP2 & RF HR 15.6 (6.7-36.4)* [C, U] *anti-CCP2 &* *anti-MCV HR 14.0(6.4-31.0)* [C, U] *anti-MCV&RF HR11.5(5.4-24.5)* [C, U] *1/2/3 auto-antibodies:* *HR 3.7(1.1-12.3) C,U, HR 15.5(5.9-14.2) C,U* *HR17.1(6.8-43.3)* [C, U] *(HRs for not achieving DFR)*
v. Steenbergen et al. (2015) *(ESPOIR, n=622) - IL2RAC*		
Teitsma et al. (2017)(U-Act-Early, n=60) *No networks in CD14+ cells could identified between DFR and nDFR.*		Teitsma et al. (2017) (U-Act-Early, n=60) *Pathways related to transcription/translation related to DFR in patients treated with MTx/TOCI and pathways related to migration of white blood cells and G-protein coupled receptors in TOZI arm and pathways involved in response to bacterial/biotic relates stimulus.*
		v. Steenbergen et al. (2015) (LEAC, n=645) - IL2RA [C] *HR 2.27(1.06-4.84)*[M]*

IMAGING BIOMARKERS

×	**Sharp v.d. Heijden score**	✓
Kuijper et al. (2016)(tREACH, n=281) [L] *OR 0.993(CI not specified)*	v.d. Woude et al. 2009) (LEAC, n=454) [C] *HR 0.95*(0.90-0.99)* [U]*	↓
v.d. Woude et al. (2012)(BeSt, n=508)[L] *OR0.98(0.94-1.02)*		
v.d. Kooij et al. (2008)(BeSt, n=508)[B] *Total SHS: nDFR 4.0(1.5-9.0), nDFR 3.3(1.0-6.9)*		
v.d. Woude et al. (2012)(Leiden EAC, n=424)[L] *OR0.97(0.93-1.01)*		
Akdemir et al. (2018)(BeSt/IMPROVED, n=133/175)[L] *OR 0.94(0.83-1.07)*		

Supplementary table S6 (Continued)

	LABORATORY BIOMARKERS	
Emery et al. (2018)(PRIZE, n=65)[L] *(no estimates specified)*		
×	**Larsen score**	✓
v.d. Woude et al. (2009)(ERAS, n=895)[C] *HR 0.94(0.88-1.00)*		
×	**Erosive at baseline**	✓
v.d. Woude et al. (2012)(LEAC, n=424)[L] *OR0.52(0.99-1.01)*		
v.d. Woude et al. (2012)(BeSt, n=508)[L] *OR0.70(0.37-1.31)*		
v.d. Kooij et al. (2008)(BeSt, n=508)[B] *Erosive (%): nDFR 72%, DFR 69%*		
×	**MRI**	✓
Burgers et al. (2018)(LEAC, n=238)[C] *BME HR 0.96(0.99-1.02)*M *Synovitis 1.04(0.95-1.15)* M *Tenosynovitis 1.03(0.95-1.11),*[M]		

Legend: All factors which were statistically tested for a potential association with achieving DFR were included in these overview, categorised by type of biomarker. Effect estimates were reported. If no regression analysis was conducted, numerical values compared between DFR and nDFR were reported.

** P<0.001, * p<0.05, DFR: DMARD-free remission, nDFR: no DMARD-free remission.

[B] Differences in baseline characteristics between DFR and non-DFR tested with t-test etc. [L] Logistic regression analysis [C] Cox regression analysis [U] Univariate, [M] Multivariate

Anti-MCV: anti-mutated citrullinated vimentin, CRP: C-reactive protein, DFR: DMARD-free remission, ESR: estimated sedimentation ratio, IL: interleukin, nDFR: no DMARD-free remission, SJC: swollen joint count, symp: symptom, HR: Hazard ratio, HLA: Human leukocyte antigen, OR: Odds ratio.

Supplementary table S7 – The selection of predictors of DFR used for figure 3

×		✓	
Age			
v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 1.01(0.98-1.03)* v.d. Woude et al. (2009) (LEAC, n=454)[C] *HR 1.02(0.99-1.03)* Kuijper et al. (2016) (tREACH, n=281)[L] *OR 0.995(CI not specified)*		no associations between age and DFR have been reported within the included articles	
Gender			
v.d. Woude et al. (2009) (LEAC, n=454)[C] *Female: HR 1.28(0.74-2.19)*		Kuijper et al. (2016) (tREACH, n=281)[L] *Female: OR 0.352*[,M]* *(CI not specified)*	↑
		v.d. Woude et al. (2012) (BeSt, n=508)[L] *Male: OR 2.3 9*[M] (1.26-4.53)*	↑
BMI			
v.d. Woude et al. (2012) (LEAC, n=424)[L] *OR 0.95(0.83-1.08)* v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 0.96(0.88-1.04)*		no associations between BMI and DFR have been reported within the included articles	
Smoking			
v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 0.69(0.36-1.33)*		v.d. Woude et al. (2009) (LEAC, n=454)[C] *HR 0.56*[U] (0.34-0.94)*	↓
(shorter) Symptom duration			
v.d. Woude et al. (2012) (BeSt, n=508) (cont)[C] *O R0.99(0.98-1.00) (cont. in weeks)* Kuijper et al. (2016) (tREACH, n=281)[L] *OR 1.00 (CI not specified)* *(measure symp duration not specified)*		v.d. Linden et al. (2010) (LEAC, n=598)[C] ≥12w symptoms vs <12w *HR 1.90*[M] (1.18 - 3.05)(not achieving DFR)*	↑
Disease activity score at baseline			
Akdemir et al. (2018) (IMPROVED, n=175)[L] *OR 0.94(0.58-1.53)* *Only IMPROVED data selected for figure*		v.d. Woude et al. (2012) (BeSt. N=508)[L] *OR 0.63*[M] (0.43-0.94)* Kuijper et al. (2016) (tREACH, n=281)[L] *OR 0.587*[M] (CI not specified)*	↓ ↓

Supplementary table S7 (Continued)

×	**Swollen Joint Count at baseline**	✓	
v.d. Woude et al. (2009) (LEAC, n=454)[C] – 44-SJC *HR 1.00(0.96-1.04)* v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 1.01(0.97-1.06)*		no associations between SJC and DFR have been reported within the included articles	
×	**HAQ**	✓	
v.d. Woude et al. (2009) (LEAC, n=454)[C] - m-HAQ *HR 1.06(0.74-1.52)*		v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 0.63(0.40-0.98)*[U]* Kuijper et al. (2016) (tREACH, n=281)[L] *OR 0.515*[U]*	↓ ↓
×	**Reuma factor**	✓	
Kuijper et al. (2016) (tREACH, n=281)[L] *OR 1.399 (CI not specified)*		v.d. Linden et al. (2009) (LEAC, n=687)[C] *HR 4.7 (2.8-8.0) *[U]* *(for not achieving DFR)* v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 0.39(0.21-0.70)*[U]*	↓ ↓
×	**Anti-CCP**	✓	
Kuijper et al. (2016) (tREACH, n=281)[L] *OR 0.636 (CI not specified)*		v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 0.20(0.10-0.39)**[M]* v.d. Linden et al. (2009) (LEAC), n=687)[C] *anti-CCP-2 HR 11.6 (5.8-23.4)[U]* *anti-CCP3 HR 6.0 (3.4-10.4)[U]* *(for not achieving DFR)*	↓ ↓
×	**CRP**	✓	
v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 1.00(0.99-1.01)*		v.d. Woude et al. (2009)(LEAC, n=454)[C] *HR 0.99*[M] (0.98-1.0)*	↓
×	**ESR**	✓	
v.d. Woude et al. (2009) (LEAC, n=454)[C] *HR 0.99(0.98-1.00)* v.d. Woude et al. (2012) (BeSt, n=508)[L] *OR 0.99 (-0.98-1.00)*		no associations between ESR and DFR have been reported within the included articles	

Supplementary table S7 (Continued)

×	**Sharp v.d. Heijden score**	✓	
Kuijper et al. (2016) (tREACH, n=281) [L] *OR 0.993(CI not specified)* v.d. Woude et al. (2012) (BeSt, n=508) [L] *OR0.98(0.94-1.02)*		v.d. Woude et al. (2009) (LEAC, n=454) [C] *HR 0.95* [U] (0.90-0.99)*	↓
×	**Erosive at baseline**	✓	
v.d. Woude et al. (2012) (LEAC, n=424) [L] *OR0.52(0.99-1.01)* v.d. Woude et al. (2012) (BeSt, n=508) [L] *OR0.70(0.37-1.31)*		no associations between erosive at baseline and DFR have been reported within the included articles	
×	**Shared Epitope**	✓	
		v.d. Woude et al. (2009) (LEAC, n=454) - HLA [C] *HR 0.46 (0.29-.75)* [U]*	↓
		v.d. Woude et al. (2012) (BeSt, n=508) [L] *OR 0.46(0.25-0.85)* [U]*	↓

Legend: Based on supplementary table S6 predictors were selected for a narrative overview (figure 3). Only high and moderate-quality studies were selected which reported on factors associated with DFR, tested by means of regression techniques. When more factors were repeatedly reported by the same study, the study including the largest study population and subsequent longest follow-up were included.

** P<0.001, * p<0.05, [B] Differences in baseline characteristics between DFR and non-DFR tested with t-test etc. [L] Logistic regression analysis [C] Cox regression analysis [U] Univariate, [M] Multivariate.

Anti-MCV: anti-mutated citrullinated vimentin, CRP: C-reactive protein, DFR: DMARD-free remission, ESR: estimated sedimentation ratio, nDFR: no DMARD-free remission, SJC: swollen joint count, symp: symptom, HR: Hazard ratio, HLA: Human leukocyte antigen, OR: Odds ratio.

LITERATURE

1. Smolen JS, Landewe R, Bijlsma Jea: EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. Ann Rheum Dis 2017, 76(6):960-977.
2. Smolen JS, Landewé RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, McInnes IB, Sepriano A, van Vollenhoven RF, de Wit M et al: EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. 2020:annrheumdis-2019-216655.
3. Schett G, Emery P, Tanaka Y, Burmester G, Pisetsky DS, Naredo E, Fautrel B, van VR: Tapering biologic and conventional DMARD therapy in rheumatoid arthritis: current evidence and future directions. Ann Rheum Dis 2016, 75(8):1428-1437.
4. Kerschbaumer A, Sepriano A, Smolen JS, van der Heijde D, Dougados M, van Vollenhoven R, McInnes IB, Bijlsma JWJ, Burmester GR, de Wit M et al: Efficacy of pharmacological treatment in rheumatoid arthritis: a systematic literature research informing the 2019 update of the EULAR recommendations for management of rheumatoid arthritis. 2020:annrheumdis-2019-216656.
5. Kuijper TM, Lamers-Karnebeek FB, Jacobs JW, Hazes JM, Luime JJ: Flare Rate in Patients with Rheumatoid Arthritis in Low Disease Activity or Remission When Tapering or Stopping Synthetic or Biologic DMARD: A Systematic Review. J Rheumatol 2015, 42(11):2012-2022.
6. O'Mahony R, Richards A, Deighton C, Scott D: Withdrawal of disease-modifying antirheumatic drugs in patients with rheumatoid arthritis: a systematic review and meta-analysis. Annals of the Rheumatic Diseases 2010, 69(10):1823-1826.
7. Ajeganova S, van Steenbergen HW, van Nies JA, Burgers LE, Huizinga TW, AH vdH-vM: Disease-modifying antirheumatic drug-free sustained remission in rheumatoid arthritis: an increasingly achievable outcome with subsidence of disease symptoms. Ann Rheum Dis 2016, 75(5):867-873.
8. Filipowicz-Sosnowska A: Drug-free remission: the goal of the future in management of patients with rheumatoid arthritis. Reumatologia 2017, 55(6):284-289.
9. Ajeganova S, Huizinga T: Sustained remission in rheumatoid arthritis: latest evidence and clinical considerations. Ther Adv Musculoskelet Dis 2017, 9(10):249-262.
10. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, Clarke M, Devereaux PJ, Kleijnen J, Moher D: The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. 2009, 339:b2700.
11. Higgins JPT GS: Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.
12. International prospective register of systematic reviews PROSPERO NIfHR: https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=132558.
13. Klarenbeek NB, van der Kooij SM, Guler-Yuksel M, van Groenendael JH, Han KH, Kerstens PJ, Huizinga TW, Dijkmans BA, Allaart CF: Discontinuing treatment in patients with rheumatoid arthritis in sustained clinical remission: exploratory analyses from the BeSt study. Ann Rheum Dis 2011, 70(2):315-319.

14. Akdemir G, Heimans L, Bergstra SA, Goekoop RJ, van OM, van Groenendael JHLM, Peeters AJ, Steup-Beekman GM, Lard LR, de Sonnaville PBJ et al: Clinical and radiological outcomes of 5-year drug-free remission-steered treatment in patients with early arthritis: IMPROVED study. Ann Rheum Dis 2018, 77(1):111-118.
15. Emery P, Burmester GR, Bykerk VP, Combe BG, Furst DE, Barre E, Karyekar CS, Wong DA, Huizinga TW: Evaluating drug-free remission with abatacept in early rheumatoid arthritis: results from the phase 3b, multicentre, randomised, active-controlled AVERT study of 24 months, with a 12-month, double-blind treatment period. Ann Rheum Dis 2015, 74(1):19-26.
16. Kuijper TM, Luime JJ, de Jong PH, Gerards AH, van ZD, Tchetverikov I, de Sonnaville PB, van Krugten MV, Grillet BA, Hazes JM et al: Tapering conventional synthetic DMARDs in patients with early arthritis in sustained remission: 2-year follow-up of the tREACH trial. Ann Rheum Dis 2016, 75(12):2119-2123.
17. Bijlsma JWJ, Welsing PMJ, Woodworth TG, Middelink LM, Petho-Schramm A, Bernasconi C, Borm MEA, Wortel CH, Ter Borg EJ, Jahangier ZN et al: Early rheumatoid arthritis treated with tocilizumab, methotrexate, or their combination (U-Act-Early): a multicentre, randomised, double-blind, double-dummy, strategy trial. Lancet 2016, 388(10042):343-355.
18. Emery P, Hammoudeh M, FitzGerald O, Combe B, Martin-Mola E, Buch MH, Krogulec M, Williams T, Gaylord S, Pedersen R et al: Sustained remission with etanercept tapering in early rheumatoid arthritis. N Engl J Med 2014, 371(19):1781-1792.
19. Kita J, Tamai M, Arima K, Nakashima Y, Suzuki T, Kawashiri SY, Iwamoto N, Okada A, Koga T, Yamasaki S et al: Treatment discontinuation in patients with very early rheumatoid arthritis in sustained simplified disease activity index remission after synthetic disease-modifying anti-rheumatic drug administration. Mod Rheumatol 2012, 22(3):346-352.
20. Huizinga TW, Conaghan PG, Martin-Mola E, Schett G, Amital H, Xavier RM, Troum O, Aassi M, Bernasconi C, Dougados M: Clinical and radiographic outcomes at 2 years and the effect of tocilizumab discontinuation following sustained remission in the second and third year of the ACT-RAY study. Ann Rheum Dis 2015, 74(1):35-43.
21. Kaneko Y, Kato M, Tanaka Y, Inoo M, Kobayashi-Haraoka H, Amano K, Miyata M, Murakawa Y, Yasuoka H, Hirata S et al: Tocilizumab discontinuation after attaining remission in patients with rheumatoid arthritis who were treated with tocilizumab alone or in combination with methotrexate: results from a prospective randomised controlled study (the second year of the SURPRISE study). Ann Rheum Dis 2018, 77(9):1268-1275.
22. Haschka J, Englbrecht M, Hueber AJ, Manger B, Kleyer A, Reiser M, Finzel S, Tony HP, Kleinert S, Feuchtenberger 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 2016, 75(1):45-51.
23. Ten WS, Hermans J, Breedveld FC, Dijkmans BA: Effect of resumption of second line drugs in patients with rheumatoid arthritis that flared up after treatment discontinuation. Ann Rheum Dis 1997, 56(4):235-239.
24. Brocq O, Millasseau E, Albert C, Grisot C, Flory P, Roux CH, Euller-Ziegler L: Effect of discontinuing TNFalpha antagonist therapy in patients with remission of rheumatoid arthritis. Joint Bone Spine 2009, 76(4):350-355.

25. Nishimoto N, Amano K, Hirabayashi Y, Horiuchi T, Ishii T, Iwahashi M, Iwamoto M, Kohsaka H, Kondo M, Matsubara T et al: Drug free REmission/low disease activity after cessation of tocilizumab (Actemra) Monotherapy (DREAM) study. Mod Rheumatol 2014, 24(1):17-25.
26. El MY, El GM, Youssef S, Ahmed I, Bahlas S, Hegazi M, Nasr A: Optimizing therapy in inflammatory arthritis: prediction of relapse after tapering or stopping treatment for rheumatoid arthritis patients achieving clinical and radiological remission. Clin Rheumatol 2016, 35(12):2915-2923.
27. Klarenbeek NB, Guler-Yuksel M, van der Kooij SM, Han KH, Ronday HK, Kerstens PJ, Seys PE, Huizinga TW, Dijkmans BA, Allaart CF: The impact of four dynamic, goal-steered treatment strategies on the 5-year outcomes of rheumatoid arthritis patients in the BeSt study. Ann Rheum Dis 2011, 70(6):1039-1046.
28. Ajeganova S, van Steenbergen HW, van Nies JAB, Burgers LE, Huizinga TWJ, Mil vdH-v: Disease-modifying antirheumatic drug-free sustained remission in rheumatoid arthritis: an increasingly achievable outcome with subsidence of disease symptoms. Archives of Disease in Childhood: Education and Practice Edition 2015(no pagination).
29. Tiippana-Kinnunen T, Paimela L, Kautiainen H, Laasonen L, Leirisalo-Repo M: Can disease-modifying anti-rheumatic drugs be discontinued in long-standing rheumatoid arthritis? A 15-year follow-up. Scand J Rheumatol 2010, 39(1):12-18.
30. van der Woude D, Young A, Jayakumar K, Mertens BJ, Toes RE, van der Heijde D, Huizinga TW, AH vdH-vM: Prevalence of and predictive factors for sustained disease-modifying antirheumatic drug-free remission in rheumatoid arthritis: results from two large early arthritis cohorts. Arthritis Rheum 2009, 60(8):2262-2271.
31. Versteeg GA, Steunebrink LMM, Vonkeman HE, Ten Klooster PM, van der Bijl AE, van de Laar MAFJ: Long-term disease and patient-reported outcomes of a continuous treat-to-target approach in patients with early rheumatoid arthritis in daily clinical practice. Clin Rheumatol 2018, 37(5):1189-1197.
32. van Nies JA, Tsonaka R, Gaujoux-Viala C, Fautrel B, AH vdH-vM: Evaluating relationships between symptom duration and persistence of rheumatoid arthritis: does a window of opportunity exist? Results on the Leiden early arthritis clinic and ESPOIR cohorts. Ann Rheum Dis 2015, 74(5):806-812.
33. van Mulligen E, de Jong PHP, Kuijper TM, van der Ven M, Appels C, Bijkerk C, Harbers JB, de Man Y, Molenaar THE, Tchetverikov I et al: Gradual tapering TNF inhibitors versus conventional synthetic DMARDs after achieving controlled disease in patients with rheumatoid arthritis: first-year results of the randomised controlled TARA study. 2019, 78(6):746-753.
34. Zeidler H: Drug-free sustained remission or spontaneous remission by natural history in rheumatoid arthritis? An unsolved question: comment on the article of Ajeganova et al. Ann Rheum Dis 2017, 76(7):e16.
35. Akdemir G, Markusse IM, Bergstra SA, Goekoop RJ, Molenaar ET, van Groenendael JHLM, Kerstens PJSM, Lems WF, Huizinga TWJ, Allaart CF: Comparison between low disease activity or DAS remission as treatment target in patients with early active rheumatoid arthritis. RMD Open 2018, 4(1):e000649.
36. Baker KF, Skelton AJ, Lendrem DW, Scadeng A, Thompson B, Pratt AG, Isaacs JD: Predicting drug-free remission in rheumatoid arthritis: A prospective interventional cohort study. Journal of autoimmunity 2019, 105:102298.

37. Boeters DM, Burgers LE, Sasso EH, Huizinga TWJ, van der Helm-van Mil AHM: ACPA-negative RA consists of subgroups: patients with high likelihood of achieving sustained DMARD-free remission can be identified by serological markers at disease presentation. Arthritis Res Ther 2019, 21(1):121.

38. de Moel EC, Rech J, Mahler M, Roth J, Vogl T, Schouffoer A, Goekoop RJ, Huizinga TWJ, Allaart CF, Toes REM et al: Circulating calprotectin (S100A8/A9) is higher in rheumatoid arthritis patients that relapse within 12 months of tapering anti-rheumatic drugs. Arthritis Res Ther 2019, 21(1):268.

CHAPTER 6

TAPERING TOWARDS DMARD-FREE REMISSION IN ESTABLISHED RHEUMATOID ARTHRITIS: TWO YEAR RESULTS OF THE TARA TRIAL

E van Mulligen, AEAM Weel, JMW Hazes, AHM van der Helm - van Mil,

PHP de Jong

Ann Rheum Dis. 2020;79:1174-1181.

ABSTRACT

Objectives To evaluate the two-year clinical effectiveness of two gradual tapering strategies. The first strategy consisted of tapering the csDMARD first (i.e. methotrexate in ~90%), followed by the TNF-inhibitor, the second strategy consisted of tapering the TNF-inhibitor first, followed by the csDMARD.

Methods This multicenter single-blinded randomized controlled trial included rheumatoid arthritis (RA) patients with well-controlled disease for ≥3 consecutive months, defined as a DAS44≤2.4 and a swollen joint count (SJC)≤1, which was achieved with a csDMARD and a TNF-inhibitor. Eligible patients were randomized into gradual tapering the csDMARD followed by the TNF-inhibitor, or vice versa. The primary outcome was the number of disease flares. Secondary outcomes were DMARD-free remission(DFR), DAS, functional ability(HAQ-DI), and radiographic progression.

Results 189 patients were randomly assigned to tapering their csDMARD (n=94) or TNF-inhibitor (n=95) first. The cumulative flare rate after 24-months was respectively 61% (95%CI, 50%-71%) and 62% (95%CI, 52%-72%). The patients who tapered their csDMARD first were more often able to go through the entire tapering protocol and reached DFR more often than the group that tapered the TNF-inhibitor first (32% versus 20%(p=0.12) and 21% versus 10%(p=0.07), respectively). Mean DAS and HAQ-DI over time, and radiographic progression did not differ between groups (p=0.45, p=0.17, p=0.8, respectively).

Conclusion The order of tapering did not affect flare rates, DAS or HAQ-DI. DFR was achievable in 15% of established RA patients, slightly more frequent in patients that first tapered csDMARDs. Because of similar effects from a clinical viewpoint, financial arguments may influence the decision to taper TNF-inhibitors first.

INTRODUCTION

In rheumatoid arthritis (RA) disease outcomes have improved tremendously in the last decades, mainly due to early initiation of therapy, a treat-to-target approach and intensive therapy with conventional synthetic disease-modifying anti-rheumatic drugs (csDMARDs) and biologicals. As a result, remission in RA occurs more frequently.[1]

If patients are successfully treated and the disease is well-controlled, the patient as well as the treating physician will explore the possibility to taper medication. Reasons for tapering medication are among others reduction in costs, patient preference, and prevention of (long-term) side effects. Tapering treatment may, however, lead to more transient or persistent disease flares with potential harmful consequences. [2-4] Previous research already showed that it is possible to taper DMARDs in RA and, therefore, current treatment recommendations advise to consider tapering therapy when RA patients are in sustained remission.[2, 5] However, there is no consensus on the best tapering strategy.

With the possibility to taper, the final step in tapering is to fully stop DMARDs. It has been suggested that sustained DMARD-free remission (DFR, which is defined as the absence of synovitis after cessation of DMARD therapy) is a preferred ultimate outcome of RA. Previous research in early RA populations showed that 10-20% of RA patients is able to achieve this outcome, [6, 7] which was independent of the chosen treatment strategy.[7] However, it is currently unknown if reaching DFR is a reachable outcome in established RA.

Therefore, the aim of this study is to evaluate the two-year clinical effectiveness of two gradual tapering strategies, namely tapering the csDMARD first followed by the TNF-inhibitor, or vice versa, in established RA patients. We will also explore the possibility to reach DFR within this population.

PATIENTS AND METHODS

Patient population

Patients studied were included in the Tapering strategies in Rheumatoid Arthritis (TARA) trial (NTR2754). Inclusion started September 2011 and ended July 2016. The TARA trial was a multicenter, single-blinded randomized trial, and was carried out in twelve rheumatology centers in the south-western part of the Netherlands.[8] Adult

RA patients with well-controlled disease, defined as a disease activity score (DAS)≤2.4 and a swollen joint count (SJC)≤1 at two consecutive time-points within a 3-month interval, using a combination of a csDMARD and TNF-inhibitor, were included. Medical ethics committees of each participating center approved the protocol and all patients gave written informed consent before inclusion.

Randomisation and blinding

Patients were randomized using minimization randomization stratified for center. Trained research nurses, blinded to the allocated tapering arm, examined patients and calculated the DAS.

Tapering schedule

Patients were randomized into tapering the csDMARD in the first year followed by tapering the TNF-inhibitor in the second year, or vice versa. The csDMARD as well as the TNF-inhibitor were gradually tapered to discontinuation in three steps. Tapering csDMARDs was realized by cutting the dosage into half, a quarter and thereafter it was stopped. TNF-inhibitors were tapered by doubling the dose interval, followed by cutting the dosage into half, and thereafter it was stopped. The total tapering schedule for each drug took 6 months, with dose adjustments every 3 months as long as there was still a well-controlled disease. At the start of the study, patients were asked to refrain from glucocorticoids (GCs). There were no restrictions on the use of non-steroidal anti-inflammatory drugs (NSAIDs) or intra-articular GC injections.

The protocol was terminated in case of a flare (DAS>2.4 and/or SJC>1). The previous effective dose was restarted and if necessary, medication was intensified further according to a treat-to-target approach, until low disease activity was reached again. After a flare, no further attempts were taken to taper medication during the remainder of the study.

Assessments and outcomes

Patients were examined at baseline and every 3 months thereafter. At each time point, the DAS, medication usage, and self-reported questionnaires were collected, except for hand and foot radiographs, which were obtained at baseline and after 1 and 2 years of follow-up. Throughout the whole study follow-up (serious) adverse event were recorded.

The primary outcome was the proportion of patients with a disease flare within the entire follow-up period of two years. Secondary endpoints were (1) the proportion of patients going through the entire tapering protocol, (2) DMARD-free remission, (3) disease activity, (4) functional ability, (5) radiographic progression, and (6) adverse events.

Disease activity was measured with the DAS. Functional ability was measured with the health assessment questionnaire disability index (HAQ-DI).[9] Higher HAQ-DI scores indicate poorer function. Radiographic progression was measured with the modified total Sharp score (mTSS).[10] Radiographs were scored chronologically by two out of three qualified assessors, who were blinded for study allocation and the identity of the patients.[11] Median mTSS are reported.[12] The weighted overall κ was 0.75 with >99% agreement. The percentage of patients with radiographic progression, defined as a change in mTSS>0.5 and >1.3 (the smallest detectable change over 2 years), are given.[12] Safety monitoring took place according to Dutch guidelines, and included laboratory tests every 3 months.[13-15] The medication was stopped or the dosage was lowered in case of adverse events related to medication use.

Statistical analysis

The TARA trial was a superiority trial, powered to detect a 20% difference in flare rates between both tapering strategies after one year of follow-up, using a significance level of α=0.05 and a power of 80%, which was previously described elsewhere.[16] For the current analysis, outcomes were calculated in an intention-to-treat analysis. Differences between groups in (1) cumulative flare rates, (2) proportion of patients going through the entire tapering protocol and (3) proportion of patients who reached DMARD-free remission were analyzed using logistic regression models. Missing data was imputed for these three analyses making use of using the last observation carried forward method. Flare-free survival was visualized with Kaplan-Meier curves, in which patients who were lost to follow-up were censored. Linear mixed models with maximum likelihood optimization were used to compare DAS and HAQ-DI over time. Statistical comparisons of outcomes were made by Student's t-test, χ^2 test or Wilcoxon rank-sum test when appropriate. All data was analyzed using STATA15. A p-value ≤0.05 was considered statistically significant.

Patient and public involvement

Patient partners are regularly consulted as advisor for all ongoing projects in the

Erasmus MC. The patient panel of the Erasmus MC consist of 15-20 patients of different age, sex and with different rheumatic diseases. Study results and study proposals are discussed on a regular basis. For the TARA study, patients were consulted for the design of the study, developing the research question and outcome parameters.

RESULTS

Patients

A total of 189 patients were randomly assigned to taper their csDMARD (n=94) or TNF-inhibitor (n=95) first (figure 1). After two years of follow-up 13 and 9 patients dropped out of the study, and complete follow-up data was obtained for 167 patients (figure 1). Patients had a median symptom duration of 6.2 years and were predominantly female (66%) with an average age of 56.6 years (table 1). Within the group who tapered the csDMARD first, 80% had DAS remission (DAS44<1.6), compared to 88% of patients who tapered their TNF-inhibitor first. Furthermore, 33% of patients in the group who tapered the csDMARD first and 37% of the group who tapered the TNF-inhibitor first were in Boolean remission, defined as TJC28≤1, SJC28≤1, CRP≤10mg/l, PGA≤10mm (0-100 mm scale) at baseline (table 1)

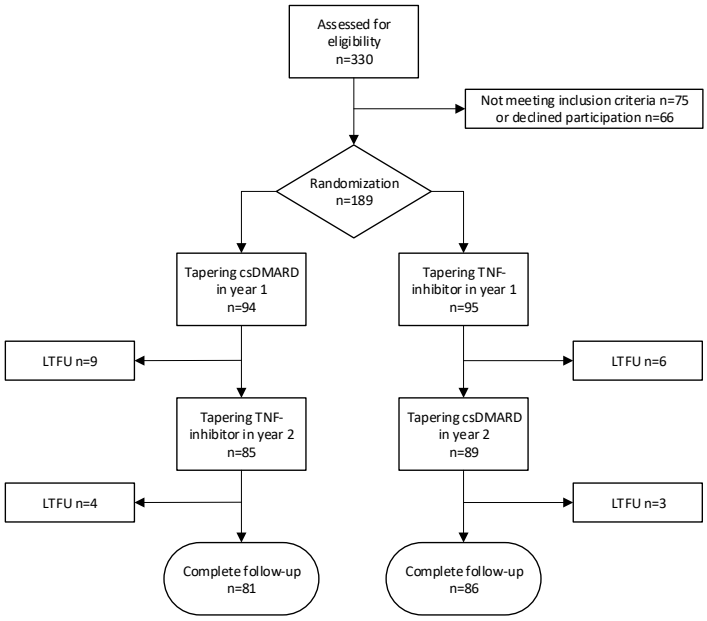


Figure 1 Flowchart of the TARA trial. Trial profile and patient participation are indicated as numbers of patients. csDMARD: conventional synthetic disease-modifying anti-rheumatic drug, LTFU: lost to follow-up, TNF-inhibitor: tumor necrosis factor inhibitor.

Table 1 Baseline characteristics of patients

Characteristics	Tapering csDMARD first (n=94)	Tapering TNF-inhibitor first (n=95)
Demographic		
Age (years), mean (95% CI)	55.9 (53.0-58.8)	57.2 (55.0-59.4)
Gender, female, n (%)	67 (71)	58 (61)
Disease characteristics		
Symptom duration (years), median (IQR)	6.0 (4.1-8.5)	6.4 (4.2-8.9)
RF positive, n (%)	50 (57)	59 (65)
ACPA positive, n (%)	62 (71)	67 (75)
Disease activity		
DAS44, mean (95% CI)	1.1 (0.9-1.2)	1.0 (0.9-1.1)
DAS clinical remission, DAS44<1.6, n (%)	75 (80)	84 (88)
Boolean remission, n (%)	31 (33)	35 (37)
HAQ-DI, mean (95% CI)	0.52 (0.42-0.62)	0.47 (0.35-0.58)
Use of csDMARDs*		
MTX monotherapy, n (%)	64 (69)	49 (52)
MTX + HCQ, n (%)	18 (19)	27 (29)
MTX + SASP + HCQ, n (%)	5 (5)	6 (6)
MTX + SASP, n (%)	3 (3)	2 (2)
MTX + LEF, n (%)	1 (1)	0 (0)
SASP monotherapy, n (%)	0 (0)	3 (3)
SASP + HCQ, n (%)	2 (2)	0 (0)
SASP + LEF, n (%)	0 (0)	1 (1)
LEF monotherapy, n (%)	1 (1)	3 (3)
LEF + HCQ, n (%)	0 (0)	1 (1)
HCQ monotherapy, n (%)	0 (0)	3 (3)
Use of TNF-inhibitor		
Etanercept, n (%)	51 (54)	52 (55)
Adalimumab, n (%)	37 (39)	40 (42)
Certolizumab, n (%)	2 (2)	2 (2)
Golimumab, n (%)	4 (4)	1 (1)
Use of glucocorticosteroids, n (%)		
Oral, n (%)	1 (1)	0 (0)
Radiographs (hand/foot)		
mTSS (0-488), median (IQR)	2 (0-6.5)	1 (0-3.5)
Erosive disease, n (%) **	37 (39)	26 (27)

*some patients used a combination of csDMARDs, ** Erosive disease is characterized as having >1 erosion in three separate joints. ACPA: anti-citrullinated protein antibody; CI: confidence interval; csDMARD: conventional synthetic disease modifying anti-rheumatic drug; DAS44: disease activity score measured in 44 joints; HAQ-DI: Health Assessment Questionnaire Disability Index; HCQ: hydroxychloroquine; IQR: interquartile range; mTSS: modified Sharp/Van der Heijde score; MTX: methotrexate; RF: rheumatoid factor; SASP: sulfasalazine.

Primary outcome

After two years of follow-up, flare rates (95% CI) were 61% (50%-71%) in the group who tapered the csDMARD first, and 62% (52%-72%) in the group who tapered the TNF-inhibitor first (p=0.84)(figure 2). The median time-to-flare (IQR) was 9.5 (6.5-21) months for patients tapering the csDMARD first, and 12 (6.5-15.5) months for patients tapering the TNF-inhibitor first. Median flare duration (IQR) was for both tapering groups 3 (3-6) months. Use of glucocorticoids was similar for both tapering arms (supplemental table S1).

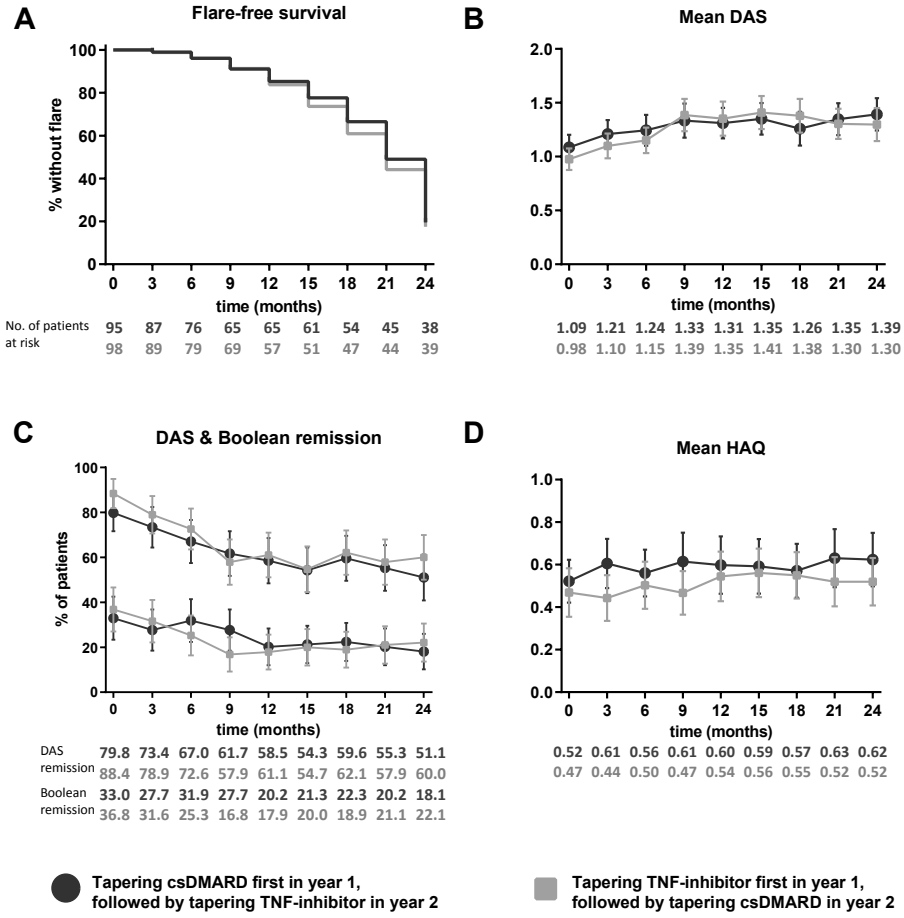


Figure 2 Disease activity over time. (A) Kaplan-Meier of flare-free survival, numbers below the graph indicate the number of patients at risk, (B) mean DAS based on 44 joints over time, (C) percentage of patients in DAS remission (DAS44<1.6) indicated with solid lines and the percentage of patients in Boolean remission: TJC28≤1, SJC28≤1, CRP≤10mg/l, PGA≤10mm (0-100 mm scale) indicated with dotted lines, (D) functional ability measured with HAQ over time. Error bars indicate 95% confidence intervals. Numbers below graphs indicate mean values of

the outcome per tapering arm, per time-point, unless other indicated. csDMARD: conventional synthetic disease-modifying anti-rheumatic drug, DAS: disease activity score, HAQ: health assessment questionnaire, TNF-inhibitor: tumor necrosis factor inhibitor.

DMARD-free remission

Of the patients who tapered respectively their csDMARD and TNF-inhibitor first 29 (31%) and 20 (21%) were able to go through the entire tapering protocol of tapering their TNF-inhibitor and 1 csDMARD (p=0.12, figure 3). None of these patients experienced a flare after withdrawal of the csDMARD and TNF-inhibitor (period between 18 and 24 months of follow-up). Although these patients went through the entire tapering protocol, not all of them were in DFR, because some were using a combination of csDMARDs at baseline (table 1) and in the protocol only one csDMARD was tapered. This means that from the total amount of patients who tapered according to protocol, not all were in DFR. In total, 19 (20%) patients tapering csDMARDs first and 10 (11%) patients tapering TNF-inhibitor first were in DFR after 24 months of follow-up (p=0.07, figure 3). In both groups, all patients reached DFR after 18 months of follow-up, and none of them used glucocorticosteroids in the period thereafter.

Disease activity, functional ability, and radiographic progression

No significant differences were found in disease activity (p=0.45) and functional ability (p=0.17) between both tapering groups over time (figure 2). The percentage of patients in Boolean remission after 1 year of follow-up decreased from 33% to 20% in the group who tapered the csDMARD first and from 37% to 18% in the group who tapered the TNF-inhibitor first, and in the second year this percentage stabilized (figure 2).Median (IQR) mTSS scores were 3 (0-7.5) in the csDMARD and 1 (0-4.5) in the TNF inhibitor tapering group after 2 year of follow-up. The cumulative probability plots of both groups were similar (figure 4). Radiographic progression, defined as an mTSS increase of >1.3, occurred in 6.1% of the patients in the csDMARD-tapering group and 7.5% of the patients in the TNF inhibitor tapering group (p=0.8). These percentage were respectively 16.3% and 20% if we use an mTSS increase of >0.5 as definition for radiographic progression (p=0.9). An increase in erosive disease (>1 erosion in 3 separate small hand or feet joints) after 2 years of follow-up was observed in 6.4% of patients who tapered the csDMARD first, and in 11.6% of patients who tapered their TNF-inhibitor first.

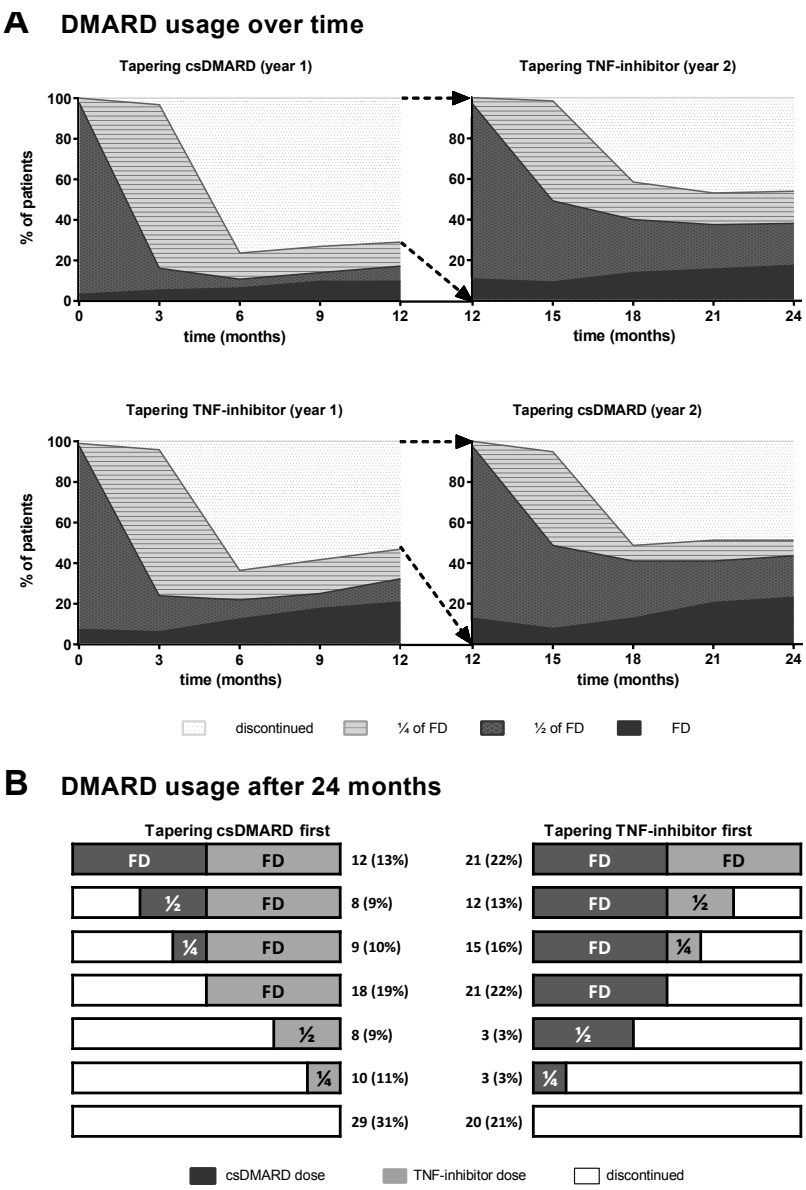


Figure 3 Overview of medication use throughout two years of follow-up. In the first year data was used of all patients, for the second year only data was shown for patients who actually tapered their medication. When patients had a flare, it was no longer allowed to continue tapering throughout the rest of the study. (A) DMARD usage over time indicated for the two tapering arms, given as percentages of patients. (B) DMARD usage after 24 months. Each bar represents a certain dosage of the csDMARD and the TNF-inhibitor, ranging from no tapering on top (full dose, FD) to discontinuation of the csDMARD and the TNF-inhibitor below. Numbers (%) next to bars indicate the number of patients who reached the indicated level of tapering after following the protocol for 24 months, as a percentage of the original TARA population. csDMARD: conventional synthetic disease-modifying anti-rheumatic drug, FD: full dose or the original dose before tapering commenced, TNF-inhibitor: tumor necrosis factor inhibitor.

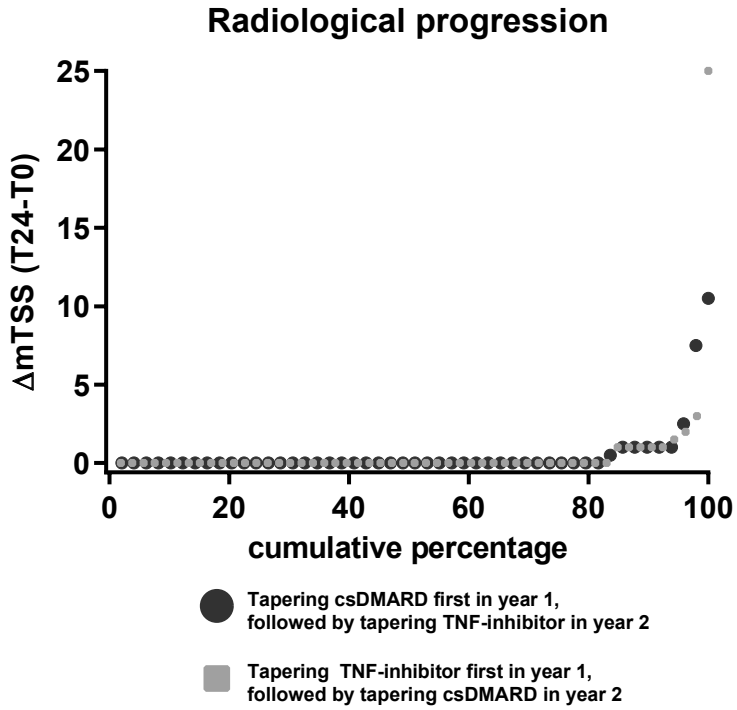


Figure 4 Radiological progression within the two years of follow-up. Radiological progression was measured with the modified total Sharp score (mTSS). csDMARD: conventional synthetic disease-modifying anti-rheumatic drug, mTSS: modified total Sharp Score, TNF-inhibitor: tumor necrosis factor inhibitor.

Adverse events

In total, 15 (8%) serious adverse events were reported. None of them were reported as being related to the study medication (table 2). At baseline, self-reported adverse events (AEs) were collected and 47.1% of all patients reported side effects. No differences were found between both tapering groups regarding the number of AEs reported and the burden of AEs (table 2). However, when assessing each drug separately then AEs were more often reported for methotrexate compared to the TNF-inhibitor (35% vs. 23%, p=0.02). The AEs related to MTX also had more impact on patients' life compared to AEs related to the TNF-inhibitor (20 vs. 8.8, p<0.0001, measured with a visual analogue scale)(table 2). The self-reported AEs and their impact on patients' lives were all measured before actual tapering commenced.

Table 2 Adverse events

Adverse events	Tapering csDMARD first (n=94)	Tapering TNF-inhibitor first (n=95)
MTX		
Patients reporting AE due to MTX	32 (34)	35 (37)
Off day	7 (7)	8 (8)
Nausea	22 (23)	18 (19)
Fatigue	7 (7)	8 (8)
Acne	0 (0)	3 (3)
Hair loss	5 (5)	5 (5)
Abnormalities of oral mucosa	1 (1)	3 (3)
Headache	1 (1)	0 (0)
Burden of AE due to MTX, VAS (0-100), mean (sd)	20 (27)	20 (27)
TNF-inhibitor		
Patients reporting AE due to TNF-inhibitor	23 (24)	21 (22)
Pain of injection	13 (14)	7 (7)
Fear of injection	3 (3)	4 (4)
Irritation at place of injection	8 (9)	10 (11)
General skin changes	5 (5)	3 (3)
Itch	1 (1)	1 (1)
Gastro-intestinal complaints	2 (2)	2 (2)
Fatigue	2 (2)	1 (1)
Burden of AE due to TNF-inhibitor, VAS (0-100), mean (sd)	7.6 (12)	10 (19)
Serious adverse events*	10 (12)	5 (6)

*Serious AEs per tapering arm were respectively: tapering csDMARDs first 7x hospitalization (3x total hip replacement surgery, 1x pneumonia, 1x decompression shoulder, 1x pancreatitis, 1x angina pectoris), 1x herpes zoster, 1x basal cell carcinoma, 1x large-cell lung carcinoma; tapering TNF-inhibitor first 4x hospitalization (2x peripheral vascular disease, 1x total knee replacement, 1x myocardial infarction), 1x bruised rib. csDMARD: conventional synthetic disease modifying anti-rheumatic drug; VAS: visual analogue scale

DISCUSSION

In this study, the two-year clinical effectiveness of two gradual tapering strategies in established RA were evaluated. The first strategy consisted of tapering the csDMARD first followed by the TNF-inhibitor, the second strategy consisted of tapering the TNF-inhibitor first, followed by the csDMARD. After two years of follow-up, 61% and 62% of patients who respectively tapered their csDMARD or TNF-inhibitor first experienced a disease flare. Also, no differences were seen in disease activity, functional ability, radiographic progression, and serious adverse events. Furthermore, 31% and 21% of patients were able to complete the entire tapering protocol. After two years, 20% and 11% of patients were in DMARD-free remission.

The flare rates within the TARA trial were high, but within the range of previous reported flare rates (51%-77%).[17-20] Also, our median flare duration, which was 3 months, is comparable with previous tapering studies.[3] This underlines the robustness of the current data and suggests that these flare rates are generalizable to clinical practice.

DFR is nowadays the closest to actual cure of RA, which might be reached by controlled tapering of medication in part of the patients. However, data on achieving DFR in established RA patients are sparse. The RETRO study showed that 13 out of 27 established RA patients (48%) were able to reach DFR. However, these data were based on a very low sample size.[21] Our DFR rate is comparable with the Leiden Early Arthritis cohort (LEAC), 158/889 (17.8%), however direct comparison is hampered due to various reasons, among which the difference in study design, disease stage (early versus established RA) and duration of being in DFR.[6] In particular, the duration of DFR is an important measure of sustainability, and inversely related to the frequency of disease flares.[22]

In both tapering groups all patients reached DFR after 18 months of follow-up. Interestingly, none of those patients experienced a flare in the 6 months after DMARD stop, whereas other studies reported flare rates between 5-25% in the first 6 months after achieving DFR.[23-26] Since clearance can take more than 6 months for certain TNF-inhibitors, we might have overestimated the proportion of patients in DFR, in the group who tapered their csDMARDs first. Nonetheless, differences between groups were not significant, and we found similar flare rates in both tapering groups, which indicates that our final results are valid. Still, optimal follow-up for assessing DFR should be longer than 6 months.

A limitation of the TARA trial is that we allowed the use of >1 csDMARD. Because only one of them was tapered according to protocol, not all patients who went through the entire tapering protocol were in DFR. Ideally, we should have included only those patients who used one csDMARD combined with a TNF-inhibitor. However, subgroup analysis revealed that tapering was not more successful in patients who used multiple csDMARDs compared to the patients who used only one csDMARD.

One could argue that tapering should only take place when patients are in a "deep" sustained remission to increase the chance at DFR and to minimize the risk of flare. Current EULAR guidelines advise to only taper medication in case of persistent remission, preferably Boolean-based.[5] For the TARA trial we used a DAS<2.4 combined with maximum of 1 swollen joint, instead of the proposed remission criteria by the EULAR. This was chosen, because of a low inclusion rate. Furthermore, at time of recruitment another trial was setup making use of the same eligibility criteria. Although we used less stringent criteria to start tapering therapy, our flare rates were comparable to other tapering strategies. Furthermore, within our study no association was found between being in Boolean remission at baseline and staying flare-free during follow-up. This suggests that Boolean remission on its own is not a good predictor for flare-free survival when medication is tapered. Moreover, if persistent Boolean remission is the prerequisite for tapering therapy fewer patients will be eligible for tapering, while in our trial only 33 (17%) patients were not able to taper any treatment.

Although 15% of our established RA patients were able to reach DFR, it is arguable whether this outweighs the risk of a disease flare (61% in our study). Especially, since it was recently shown that disease flares have a significant effect on patients' lives, with a duration of more than 6 months.[27, 28]

Ideally, rheumatologists want to be more certain about which patient is able to taper successfully, as current tapering strategies are based upon a trial-and-error approach, which results in high flare rates that significantly influence patients' lives. Unfortunately, we still do not know which patients are more eligible for tapering and whom will have a higher chance at reaching DFR. Present data (re)confirmed that tapering treatment is possible and that DFR is achievable in a small proportion of patients even within those with an established RA. In our opinion, future studies should focus on patient subsets eligible to (continue) taper medication to reduce the amount of flares and to increase the number of patients that reach DFR.

In conclusion, the order of tapering did not affect flare rates, disease activity or physical functioning. In total, 61% of patients had a flare in the two years of follow-up. DFR was achievable in a small proportion of patients and was seen slightly more frequent in patients that tapered their csDMARDs first. Because of similar effects from a clinical perspective, financial arguments may influence the decision to taper TNF-inhibitors first.

Supplemental table S1 Glucocorticoid use within two years of follow-up

	Tapering csDMARD first (n=94)	**Tapering TNF-inhibitor first (n=95)**
Oral glucocorticoids*, n (%)	5 (5)	6 (6)
Intra-articular glucocorticoids*, n (%)	20 (21)	23 (24)
Intramuscular glucocorticoids*, n (%)	7 (7)	11 (12)

*Numbers indicate number of patients who used glucocorticoids within the two years of follow-up.

csDMARD: conventional synthetic disease modifying anti-rheumatic drug, TNF: tumor necrosis factor.

LITERATURE

1. Aletaha D, Smolen JS: Diagnosis and Management of Rheumatoid Arthritis: A Review. *Jama* 2018, 320(13):1360-1372.
2. Schett G, Emery P, Tanaka Y, Burmester G, Pisetsky DS, Naredo E, Fautrel B, van Vollenhoven R: Tapering biologic and conventional DMARD therapy in rheumatoid arthritis: current evidence and future directions. *Ann Rheum Dis* 2016, 75(8):1428-1437.
3. Kuijper TM, Lamers-Karnebeek FB, Jacobs JW, Hazes JM, Luime JJ: Flare Rate in Patients with Rheumatoid Arthritis in Low Disease Activity or Remission When Tapering or Stopping Synthetic or Biologic DMARD: A Systematic Review. *J Rheumatol* 2015, 42(11):2012-2022.
4. Henaux S, Ruyssen-Witrand A, Cantagrel A, Barnetche T, Fautrel B, Filippi N, Lukas C, Raffeiner B, Rossini M, Degboe Y *et al*: Risk of losing remission, low disease activity or radiographic progression in case of bDMARD discontinuation or tapering in rheumatoid arthritis: systematic analysis of the literature and meta-analysis. *Ann Rheum Dis* 2018, 77(4):515-522.
5. Smolen JS, Landewe RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, McInnes IB, Sepriano A, van Vollenhoven RF, de Wit M *et al*: EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. *Ann Rheum Dis* 2020.
6. Ajeganova S, van Steenbergen HW, van Nies JA, Burgers LE, Huizinga TW, van der Helm-van Mil AH: Disease-modifying antirheumatic drug-free sustained remission in rheumatoid arthritis: an increasingly achievable outcome with subsidence of disease symptoms. *Ann Rheum Dis* 2016, 75(5):867-873.
7. Akdemir G, Heimans L, Bergstra SA, Goekoop RJ, van Oosterhout M, van Groenendael J, Peeters AJ, Steup-Beekman GM, Lard LR, de Sonnaville PBJ *et al*: Clinical and radiological outcomes of 5-year drug-free remission-steered treatment in patients with early arthritis: IMPROVED study. *Ann Rheum Dis* 2018, 77(1):111-118.
8. Trial NL2625 (NTR2754) [trialregister.nl/trial/2625]
9. Siegert CE, Vleming LJ, Vandenbroucke JP, Cats A: Measurement of disability in Dutch rheumatoid arthritis patients. *Clin Rheumatol* 1984, 3(3):305-309.
10. van der Heijde D: How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol* 2000, 27(1):261-263.
11. van Tuyl LH, van der Heijde D, Knol DL, Boers M: Chronological reading of radiographs in rheumatoid arthritis increases efficiency and does not lead to bias. *Ann Rheum Dis* 2014, 73(2):391-395.
12. van der Heijde D, Simon L, Smolen J, Strand V, Sharp J, Boers M, Breedveld F, Weisman M, Weinblatt M, Rau R *et al*: How to report radiographic data in randomized clinical trials in rheumatoid arthritis: guidelines from a roundtable discussion. *Arthritis Rheum* 2002, 47(2):215-218.
13. NVR Richtlijnen medicijnen: biologicals [https://www.nvr.nl/wp-content/uploads/2014/11/NVR-Medicijnen-Update_Biologicals_richtlijn-23-6-2014.pdf]
14. NVR Richtlijnen medicijnen: sulfazalazine [https://www.nvr.nl/wp-content/uploads/2014/11/NVR-Medicijnen-richtlijn-Sulfasalazine-2002.pdf]

15. NVR Richtlijnen medicijnen: methotrexaat [https://www.nvr.nl/wp-content/uploads/2014/11/NVR-Medicijnen-MTX-richtlijn-2009-update-2011.pdf]

16. van Mulligen E, de Jong PHP, Kuijper TM, van der Ven M, Appels C, Bijkerk C, Harbers JB, de Man Y, Molenaar THE, Tchetverikov I *et al*: Gradual tapering TNF inhibitors versus conventional synthetic DMARDs after achieving controlled disease in patients with rheumatoid arthritis: first-year results of the randomised controlled TARA study. *Annals of the Rheumatic Diseases* 2019:annrheumdis-2018-214970.

17. Ghiti Moghadam M, Vonkeman HE, Ten Klooster PM, Tekstra J, van Schaardenburg D, Starmans-Kool M, Brouwer E, Bos R, Lems WF, Colin EM *et al*: Stopping Tumor Necrosis Factor Inhibitor Treatment in Patients With Established Rheumatoid Arthritis in Remission or With Stable Low Disease Activity: A Pragmatic Multicenter, Open-Label Randomized Controlled Trial. *Arthritis Rheumatol* 2016, 68(8):1810-1817.

18. van Herwaarden N, van der Maas A, Minten MJ, van den Hoogen FH, Kievit W, van Vollenhoven RF, Bijlsma JW, van den Bemt BJ, den Broeder AA: Disease activity guided dose reduction and withdrawal of adalimumab or etanercept compared with usual care in rheumatoid arthritis: open label, randomised controlled, non-inferiority trial. *Bmj* 2015, 350:h1389.

19. Fautrel B, Pham T, Alfaiate T, Gandjbakhch F, Foltz V, Morel J, Dernis E, Gaudin P, Brocq O, Solau-Gervais E *et al*: Step-down strategy of spacing TNF-blocker injections for established rheumatoid arthritis in remission: results of the multicentre non-inferiority randomised open-label controlled trial (STRASS: Spacing of TNF-blocker injections in Rheumatoid ArthritiS Study). *Ann Rheum Dis* 2016, 75(1):59-67.

20. Smolen JS, Nash P, Durez P, Hall S, Ilivanova E, Irazoque-Palazuelos F, Miranda P, Park MC, Pavelka K, Pedersen R *et al*: Maintenance, reduction, or withdrawal of etanercept after treatment with etanercept and methotrexate in patients with moderate rheumatoid arthritis (PRESERVE): a randomised controlled trial. *Lancet* 2013, 381(9870):918-929.

21. Haschka J, Englbrecht M, Hueber AJ, Manger B, Kleyer A, Reiser M, Finzel S, Tony H-P, Kleinert S, Feuchtenberger 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* 2016, 75(1):45-51.

22. Verstappen M, van Mulligen E, de Jong PHP, van der Helm-Van Mil AHM: DMARD-free remission as novel treatment target in rheumatoid arthritis: A systematic literature review of achievability and sustainability. *RMD Open* 2020, 6(1):e001220.

23. Kuijper TM, Luime JJ, de Jong PH, Gerards AH, van Zeben D, Tchetverikov I, de Sonnaville PB, van Krugten MV, Grillet BA, Hazes JM *et al*: Tapering conventional synthetic DMARDs in patients with early arthritis in sustained remission: 2-year follow-up of the tREACH trial. *Ann Rheum Dis* 2016, 75(12):2119-2123.

24. Emery P, Burmester GR, Bykerk VP, Combe BG, Furst DE, Barre E, Karyekar CS, Wong DA, Huizinga TW: Evaluating drug-free remission with abatacept in early rheumatoid arthritis: results from the phase 3b, multicentre, randomised, active-controlled AVERT study of 24 months, with a 12-month, double-blind treatment period. *Ann Rheum Dis* 2015, 74(1):19-26.

25. Bijlsma JWJ, Welsing PMJ, Woodworth TG, Middelink LM, Petho-Schramm A,

Bernasconi C, Borm MEA, Wortel CH, Ter Borg EJ, Jahangier ZN *et al*: Early rheumatoid arthritis treated with tocilizumab, methotrexate, or their combination (U-Act-Early): a multicentre, randomised, double-blind, double-dummy, strategy trial. *Lancet* 2016, 388(10042):343-355.

26. Klarenbeek NB, van der Kooij SM, Guler-Yuksel M, van Groenendael JH, Han KH, Kerstens PJ, Huizinga TW, Dijkmans BA, Allaart CF: Discontinuing treatment in patients with rheumatoid arthritis in sustained clinical remission: exploratory analyses from the BeSt study. *Ann Rheum Dis* 2011, 70(2):315-319.
27. Ghiti Moghadam M, Ten Klooster PM, Vonkeman HE, Kneepkens EL, Klaasen R, Stolk JN, Tchetverikov I, Vreugdenhil SA, van Woerkom JM, Goekoop-Ruiterman YPM *et al*: Impact of Stopping Tumor Necrosis Factor Inhibitors on Rheumatoid Arthritis Patients' Burden of Disease. *Arthritis Care Res (Hoboken)* 2018, 70(4):516-524.
28. van Mulligen E, Weel AEAM, Kuijper TM, Hazes JMW, van der Helm- van Mil AHM, de Jong PHP: The impact of a disease flare during tapering of DMARDs on the lives of rheumatoid arthritis patients. *Semin Arthritis Rheum* 2020, 2020; in press.

CHAPTER 7

RESPONSE TO "TAPERING TOWARDS DMARD-FREE REMISSION IN ESTABLISHED RHEUMATOID ARTHRITIS: TWO YEAR RESULTS OF THE TARA TRIAL" BY HAROON ET AL.

E van Mulligen, P.H.P. de Jong

Ann Rheum Dis 2020; Published Online First: 14 August 2020.

We appreciate the interest in our paper by Haroon, *et al.* We presented the two year results of the TARA trial, in which we concluded that "financial arguments may influence the decision to taper TNF-inhibitors first".[1] Based on this conclusion, Haroon, *et al*. decided to respond to that with their real-world data from a resource-poor country.[2]

Ideally, if rheumatoid arthritis (RA) patients are in sustained remission, medication is quickly tapered and possibly stopped to reduce health care costs. DMARD-free remission is suggested as a preferred ultimate target in a treat-to-target management approach, however we previously showed, in a systematic literature review, that this outcome is achievable in 10-20% of the RA population.[3] Within the TARA trial we showed that DMARD-free remission was achievable in 15% of the included established RA patients. Haroon *et al.* on the other hand now report that 5 out of 45 (11%) RA and spondyloarthritis (SpA) patients were able to completely stop their bDMARDs. This confirms that DMARD-free remission is reachable for a minority of patients.

Although DMARD-free remission occurs less frequent, most of the RA patients with a well-controlled disease can lower their DMARD dosage. To illustrate, 83% of the TARA patients were able to reduce their medication dosage, which is similar to the real-world data of Haroon *et al*. Another benefit of gradual tapering with a treat-to-target approach, which includes close monitoring, is that (severe) disease flares could possibly be prevented due to slower tapering and earlier detection. In our opinion, aforementioned approach is currently the best way to taper treatment. Especially, since we have previously shown that a disease flare has a significant impact on patients' lives, which outlast the effect of a flare on disease activity.[4] Noteworthy, is the fact that although most patients reach low disease activity within 6 months after a flare, most of them have a higher disease activity post-flare compared to pre-flare.[4]

Unfortunately, current tapering strategies are still based on a trial-and-error approach which leads to high flare rates and, therefore, a tailor-made tapering approach is preferred. Moreover, no consensus had been reached on how to taper medication, because cohorts/trials directly comparing different tapering strategies are sparse.[5] Haroon, *et al.* showed that 60% of RA patients were able to reduce their bDMARD dosage when a 2-step tapering protocol was used, consisting of dose reductions every 4 months of 30% followed by 50%. Comparing this with our results from the TARA trial, in which we showed that 83% of the patients were able to reduce their

DMARD dosages with 50% every 3 months, leads to our advice to gradually taper DMARDs with 30-50% every 3-4 months in RA patients with well-controlled disease.

To summarize, by using a gradual tapering approach, almost all RA patients with a well-controlled disease can reduce their DMARD dosages. The real-world data of Haroon *et al.* underlines the fact that the majority of RA patients are able to gradually taper DMARDs.

LITERATURE

1. van Mulligen E, Weel AE, Hazes JM, et al. Tapering towards DMARD-free remission in established rheumatoid arthritis: 2-year results of the TARA trial. *Ann Rheum Dis* 2020;
2. Haroon M, Khan Z, Aamer M. Tapering Anti Rheumatic Drugs in A Resource-Poor Setting: Real World Evidence. *Ann Rheum Dis* 2020;
3. Verstappen M, van Mulligen E, de Jong PHP, et al. DMARD-free remission as novel treatment target in rheumatoid arthritis: A systematic literature review of achievability and sustainability. *RMD Open* 2020;6:e001220.
4. van Mulligen E, Weel A, Kuijper TM, et al. The impact of a disease flare during tapering of DMARDs on the lives of rheumatoid arthritis patients. *Semin Arthritis Rheum* 2020;
5. Smolen JS, Landewe RBM, Bijlsma JWJ, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. *Ann Rheum Dis* 2020;

CHAPTER 8

THE IMPACT OF A DISEASE FLARE DURING TAPERING OF DMARDS ON THE LIVES OF RHEUMATOID ARTHRITIS PATIENTS

E van Mulligen, AEAM Weel, TM Kuijper, JMW Hazes, AHM van der Helm - van Mil, PHP de Jong

Semin Arthritis Rheum. 2020;50(3):423-31

ABSTRACT

Objectives To determine the impact of a disease flare on patient reported outcome measures (PROMs) in rheumatoid arthritis (RA) patients, who are tapering treatment.

Methods Data were used from the TARA trial; a multicenter, randomized controlled trial in which RA patients, with a well-controlled disease (DAS≤2.4&SJC≤1) for at least 6 months, gradually tapered their DMARDs. PROMs of patients with a flare (DAS>2.4 and/or SJC>1) were compared every three months before and after a flare with their own norm values. Linear Mixed Models were used to investigate whether a disease flare influenced functional ability (HAQ-DI), fatigue (BRAF-MDQ), quality of life (EQ-5D and SF36), anxiety and depression (HADS), morning stiffness, general health (GH) and worker productivity, and if so, the duration was determined. For unemployment and sick leave we used descriptive statistics.

Results A flare negatively influenced GH, morning stiffness, HAQ-DI, EQ-5D, BRAF-MDQ, and the SF36 physical component scale and this effect lasted >3 months. Except for the HAQ-DI, effect sizes exceeded the minimum clinically important differences(MCIDs). For the physical outcomes effects lasted >6 months. Worker productivity was not significantly affected by a flare.

Conclusion A disease flare influenced patients' lives, the largest effect was seen in the physical outcomes, and lasted 6 months. Although on a group level effect sizes for the separate PROMs were not always significant or larger than specific MCIDs, a disease flare can still be of great importance for individual patients.

INTRODUCTION

Over the years the treatment of Rheumatoid Arthritis(RA) has improved enormously, which resulted in better outcomes, including achievement of sustained remission [1, 2]. Nowadays, 50-60% of RA patients achieve sustained remission [3, 4]. Therefore, current guidelines recommend to consider tapering treatment if patients are in sustained remission [5, 6].

Previous studies have shown that it is possible to taper biologicals, but this is accompanied with a higher chance of disease flares [7-10]. Flare rates within these studies varied from 38% to 76.6%. It has also been shown that only 41 – 67% of the patients that experienced a flare will regain remission within 6 months after treatment intensification [7, 11, 12]. Thus, many patients will have a reduced or no response to previous effective therapy, which may lead to an altered disease state or prolonged flare duration. Despite the high flare rates, current guidelines recommend to taper biologicals, which is based on a clinical and societal viewpoint.

At present, a paradigm shift in the delivery of health care is emerging, and is shifting towards patient centered healthcare. Patient-centered healthcare focuses on the individual patient preferences and needs, which can be objectified with patient reported outcome measures(PROMs) [13, 14]. In order to optimize the delivery of care during tapering we need to know how a disease flare affects these PROMs. However, data on the feasibility of tapering DMARDs from a patient's perspective are sparse.

Therefore, our objectives are (1) to determine the impact of a disease flare on patient's lives by quantifying the changes in functional ability, general health, morning stiffness, fatigue, quality of life, and worker productivity, and (2) to explore the duration of this effect.

METHODS

Study design

Data were used from the Tapering strategies in Rheumatoid Arthritis (TARA) trial (NTR2754). Adult patients with well-controlled RA, defined as a disease activity score (DAS44)≤2.4 and a swollen joint count (SJC)≤1 for at least 6 months, who were using a combination of a conventional synthetic disease modifying anti-rheumatic

drug (csDMARD) and a TNF-inhibitor, were included. Patients were randomized into gradually tapering the csDMARD or TNF-inhibitor first. In the second year, the other drug was gradually tapered. The protocol was terminated if patients experienced a flare (DAS>2.4 and/or SJC>1). The previous effective dose was restarted and if necessary, medication was intensified further according to a treat-to-target approach, until low disease activity was reached. After a disease flare it was not allowed to restart tapering [12, 15].

For the current study we compared the PROMs and DAS44, within all patients that experienced a flare, at the moment of flare, 3 months prior to a flare, and every 3 months thereafter with their own norm values. The norm was set at the average of DAS44 and PROMs 12, 9 and 6 months prior to a flare, which in our opinion was the best reference for well-controlled disease (Figure 1).

We also performed a sensitivity analysis with different flare criteria from other studies, which are less strict than our criteria, in order to assess the impact of different criteria on measured outcomes. For example, we could have classified someone as having a disease flare, while in other studies these patients would continue tapering.

Outcomes

Outcomes for the impact of a disease flare on patients' lives were DAS, general health(GH), severity of morning stiffness, functional ability, quality of life, health status, fatigue, anxiety and depression, and worker productivity.

Every three months the DAS44 and self-reported questionnaires were collected [12]. The DAS44 was used for measuring disease activity based on 44 joints[16]. The minimum clinically important difference(MCID) of the DAS44 is 0.6 [17]. GH was measured on a 0-100 mm visual analogue scale, in which 0 represented the lowest possible health state, and 100 perfect health. The MCID for GH is 10 [17]. Functional ability was measured with the health assessment questionnaire disability index (HAQ-DI)[18]. Higher scores reflect greater disability, and the MCID is 0.22 [19]. Severity of morning stiffness was measured on a 0-10 likert-scale, in which 0 represented no morning stiffness, and 10 severe morning stiffness. The MCID for morning stiffness is 1 [17]. Quality of life(QoL) was measured with the European Quality of Life – 5 Dimensions (EQ-5D). Higher scores indicate a higher quality of life, and the MCID is 0.04 [20]. Health status was measured with the short form 36 (SF36), the higher

the score, the better the health status [21-23]. The MCID of the SF36 is between 3 and 5 [23]. Fatigue was measured with the Bristol Rheumatoid Arthritis Fatigue Multi-dimensional Questionnaire (BRAF-MDQ). Higher scores represent higher levels of fatigue [24]. The MCID is 2.6 [25]. Anxiety and depression were measured with the Hospital Anxiety and Depression Scale (HADS), in which higher scores represent more anxiety and/or depression [26]. The HADS MCID for RA patients is unknown, however other chronic diseases show an MCID of 1.7 [27, 28]. Worker productivity was assessed with the iMTA Productivity cost Questionnaire (iPCQ) that addressed sick leave, reduction in work time, and productivity loss [29]. For all outcomes, the effect sizes were compared to aforementioned MCIDs.

Statistical analysis

We used data from patients that experienced a flare to determine the impact and duration of a flare on DAS44 and PROMs. The moment of flare was set as T0 and we only took the first flare into account. We used Linear Mixed Models (LMMs) with a random intercept and an autoregressive covariance matrix, to account for repeated measurements within individuals, to compare DAS44 and PROMs 3 months prior to a flare, at the moment of flare, and 3, 6, 9, and 12 months after a flare with norm values. For each patient the norm was set at the average value of DAS44 and PROMs for the combined values obtained at 12, 9 and 6 months prior to a flare. This was based on the mean DAS44 graph that showed minimal fluctuations between aforementioned timepoints in patients who experienced a flare and at those time-points these patients still had a well-controlled disease(Figure 1). Because of aforementioned reasoning we had to exclude 17, because they experienced a flare within the first 3 months of follow-up and, therefore, we could not set a norm value for these patients.

First, we examined whether there was a difference in each PROM and DAS44 over time. If there was a significant difference, the duration of this effect was determined. The duration was calculated by comparing each time-point separately with the norm, using aforementioned LMMs. For worker productivity we used descriptive statistics.

For visualization purposes, we also plotted the patients that did not have a disease flare. In this group we reclassified the 12 month visit as the new T0, because mean (sd) time to flare was 12 (6.7) months.

Outcomes were calculated in an intention-to-treat analysis, using all available data. A Bonferroni correction was applied to account for multiple testing. The calculated p-values for the impact of a flare on PROMs or DAS44 were corrected by multiplying the p-value with the total number of variables tested (n=11). The calculated p-values for the duration of a disease flare were multiplied with the total number of measurements tested (n=42). In this manner we could still consider a p-value ≤0.05 statistically significant. Corrected and uncorrected p-values are reported. All data were analyzed using STATA 15.

RESULTS

Patients

A total of 189 patients were randomized, of those 113 patients experienced a flare. Table 1 shows the norm values for patients with and without a flare. Disease characteristics and PROMs were the same for both groups, except for DAS44 (sd), which was 0.86 (0.50) in the non-flare group and 1.08 (0.52) in the flare group (p=0.0055). This difference is probably caused by a significant difference in Erythrocyte Sedimentation Rate(ESR) between both groups(p=0.008).

Table 1 Patient characteristics

Characteristics	Patients with flare (n=113)	Patients without flare (n=76)	P-value
Demographic at moment of randomization			
• Age (years), mean (sd)	58.2 (12.0)	54.1 (12.8)	0.025
• Gender, female, n (%)	77 (68.1)	48 (63.2)	0.48
Disease characteristics at moment of randomization			
• Symptom duration (years), median (IQR)	6.1 (4.3-9.1)	6.2 (3.8-8.5)	0.42
• RF positive, n (%)	61 (58.7)	45 (63.4)	0.53
• ACPA positive, n (%)	75 (72.8)	52 (74.3)	0.83
Treatment at moment of randomization			
• MTX, n (%)	106 (94)	68 (89)	0.28
• Anti-TNF, n (%)			
- Etanercept	65 (58)	38 (50)	0.31
- Adalimumab	43 (38)	33 (43)	0.46

Table 1 Patient characteristics (Continued)

Characteristics	Patients with flare (n=113)	Patients without flare (n=76)	P-value
Norm values			
Disease activity			
• DAS44, mean (sd)	1.08 (0.52)	0.86 (0.50)	0.0055
• TJC44, median (IQR)	0 (0-1)	0 (0-0)	0.16
• SJC44, median (IQR)	0 (0-0)	0 (0-0)	0.12
• General health (0-100 mm), median (IQR)	14 (5-27)	14 (2-25.5)	0.82
• ESR (mm/h), median (IQR)	9.5 (5-16)	6 (2-12)	0.008
• CRP (mg/L), median (IQR)	2 (1-5)	2 (1-5.2)	0.73
• Erosive disease on initial radiograph, n (%)a	40 (42)	29 (38)	0.64
• Morning stiffness, severity 0-10, median (IQR)	1 (0-3)	1 (0-4)	0.46
Patient reported outcomes			
• HAQ, median (IQR)	0.38 (0.13-0.75)	0.25 (0-0.63)	0.57
• EQ-5D index, mean (sd)	0.86 (0.12)	0.87 (0.12)	0.51
• BRAF-MDQ, mean (sd)	16.2 (11)	16.6 (12)	0.80
• SF36, mean (sd)			
- PCS	42.1 (11)	41.6 (11)	0.79
- MCS	56.6 (10)	56.4 (9.0)	0.91
• HADS, mean (sd)			
- Anxiety	3.6 (3.0)	4.0 (2.8)	0.39
- Depression	2.0 (1.9)	2.7 (3.0)	0.08
• Worker productivity (0-10), median (IQR)	8 (6-10)	8 (6-9)	0.14

[a]Erosive disease is characterized as having >1 erosion in three separate joints. ACPA: anti-citrullinated protein antibody; CRP: C-reactive protein; csDMARDs: conventional synthetic DMARD; DAS: disease activity score; ESR: erythrocyte sedimentation rate; EQ5D: European Quality of Life – 5 Dimensions; HAQ: Health Assessment Questionnaire; IQR: inter quartile range; MCS: mental component scale; PCS: physical component scale; RF: rheumatoid factor; sd: Standard Deviation; SJC: swollen joint count; TJC: tender joint count.

Clinical outcomes

At the moment of flare(DAS44>2.4 or SJC>1), mean DAS44 (sd) was higher in the flare group(1.84 [0.76]) compared to the non-flare group (1.04 [0.51])(figure 1A). Most of the separate components of the DAS44; TJC44, SJC44, and general health (GH); were also higher in the flare group(figure 1B, C, E). We found an overall significant effect for the DAS44 compared to the norm($p<0.0001$, table 2). The same accounted

for the DAS44 components, namely GH($p<0.0001$, table 2), SJC44($p<0.0001$), TJC44($p<0.0001$), ESR($p<0.0001$), and CRP ($p<0.0007$)(data not shown). The effect of a flare on DAS44 and GH lasted >12 months, while the clinically meaningful effect lasted 6 months (MCID DAS44>0.6 and MCID GH>10)[17].

The degree of morning stiffness, ranging from 0-10, was on average 3.7 (sd 2.8) in the flare group, and 2.5(sd 2.3) in the non-flare group at T0 (figure 2B). The degree of morning stiffness significantly differed over time($p<0.0001$, table 2). When comparing the separate time-points to the norm, we found that morning stiffness significantly worsens at the moment of flare and regains its norm value 9 months after a flare. At the moment of flare and 3 months thereafter the difference with the norm was also above the MCID of 1 (table 3)[17].

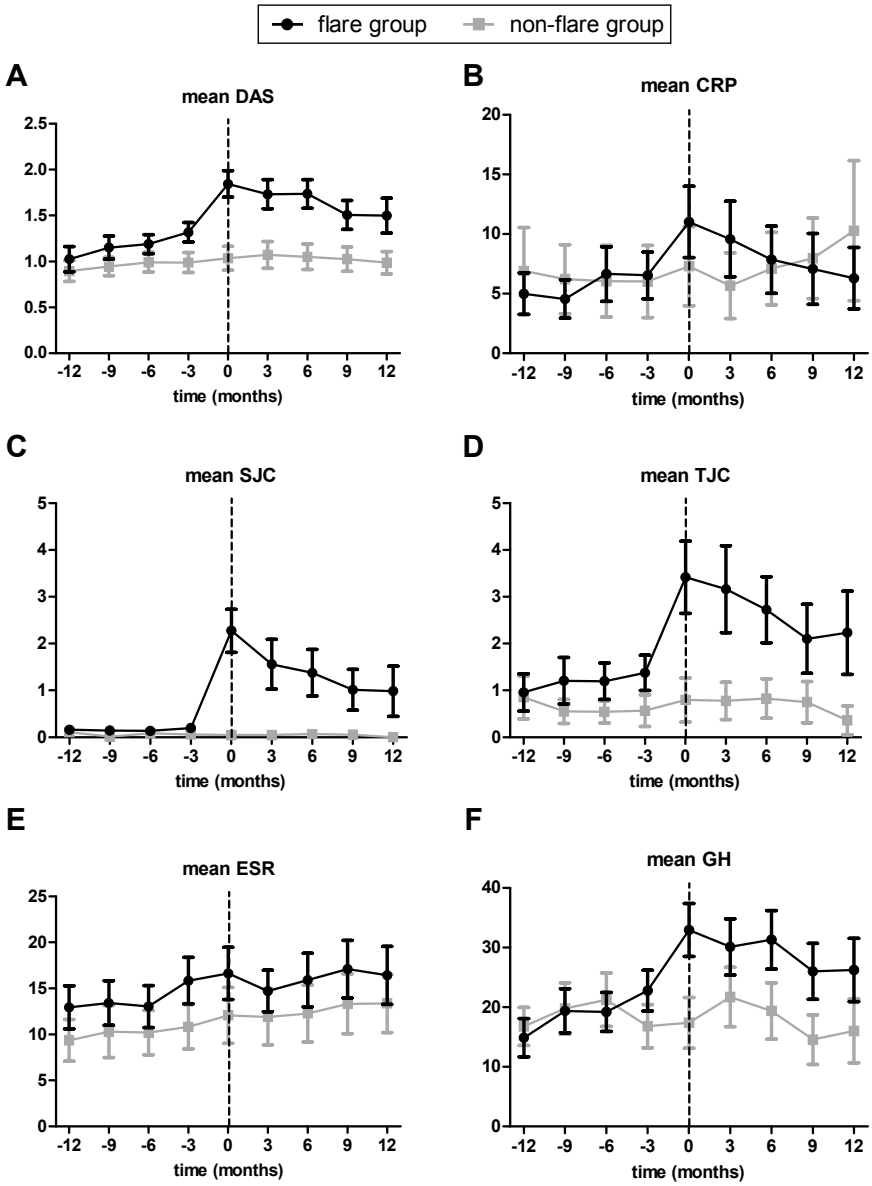


Figure 1 Clinical outcomes. (A) DAS44 scores for the flare group and the non-flare group with corrected time-points. (B) mean CRP, (C-F) separate components of the DAS44 scores: mean swollen joint count in 44 joints (SJC44), mean tender joint count in 44 joints (TJC44), mean erythrocyte sedimentation rate (ESR), and visual analogue scale for general health (GH).

Table 2 Overall differences between norm and moments thereafter.

Patient reported outcomes	P-value	Bonferroni corrected p-value[a]
DAS44	P<0.0001	P<0.0001
VAS general health	P<0.0001	P<0.0001
Morning stiffness	P<0.0001	P<0.0001
HAQ-DI	P<0.0001	P=0.0003
SF36 PCS	P<0.0001	P=0.0004
SF36 MCS	P=0.68	P=1
EQ5D	P<0.0001	P<0.0001
BRAF-MDQ	P=0.0037	P=0.041
HADS anxiety	P=0.75	P=1
HADS depression	P=0.62	P=1
Worker productivity	P=0.32	P=1

[a]n=11. BRAF-MDQ: Bristol Rheumatoid Arthritis fatigue multidimensional questionnaire; DAS: disease activity score; EQ5D: European quality of life with 5 dimensions; HADS: hospital anxiety and depression scale; HAQ-DI: health assessment questionnaire disability index; MCS: mental component scale; PCS: physical component scale; SF36: short form 36; VAS: visual analogue scale.

Table 3 Comparison of separate time-points with the norm values to assess the duration of the effect of flare.

		Difference with norm (effect size)	95% CI			P-value	Bonferroni corrected p-value[a]
DAS44 (MCID=0.6) [17]	-T3	0.16	0.039	-	0.27	0.0089	0.37
	T0	0.68	0.56	-	0.81	<0.0001	<0.0001
	T3	0.57	0.44	-	0.70	<0.0001	<0.0001
	T6	0.57	0.43	-	0.71	<0.0001	<0.0001
	T9	0.33	0.18	-	0.47	<0.0001	0.0004
	T12	0.32	0.16	-	0.47	0.0001	0.0027
General health (MCID=10) [17]	-T3	4.4	1.10	-	7.78	0.0091	0.38
	T0	14.8	11.18	-	18.35	<0.0001	<0.0001
	T3	12.4	8.53	-	16.21	<0.0001	<0.0001
	T6	12.7	8.65	-	16.74	<0.0001	<0.0001
	T9	7.6	3.34	-	11.92	0.0005	0.021
	T12	7.8	3.22	-	12.48	0.0009	0.037
Morning stiffness (MCID=1) [17]	-T3	0.41	0.028	-	0.78	0.036	1
	T0	1.32	0.93	-	1.72	<0.0001	<0.0001
	T3	1.15	0.73	-	1.57	<0.0001	<0.0001
	T6	0.86	0.41	-	1.30	0.0001	0.0062
	T9	0.26	-0.21	-	0.73	0.28	1
	T12	0.87	0.36	-	1.37	0.0007	0.031
HAQ-DI (MCID=0.22) [19]	-T3	0.016	-0.039	-	0.071	0.57	1
	T0	0.13	0.074	-	0 .19	<0.0001	0.0002
	T3	0.12	0.065	-	0.18	<0.0001	0.0019

Table 3 (Continued)

		Difference with norm (effect size)	95% CI			P-value	Bonferroni corrected p-value[a]
	T6	0.078	0.015	-	0.14	0.015	0.61
	T9	0.046	- 0.021	-	0.11	0.18	1
	T12	0.083	0.011	-	0.16	0.025	1
SF36 PCS (MCID=3-5) [23]	-T3	-1.03	-3.67	-	1.62	0.45	1
	T0	-4.25	-6.57	-	-1.93	0.0003	0.014
	T3	-4.05	-6.84	-	-1.27	0.0044	0.18
	T6	-3.95	-6.64	-	-1.26	0.0041	0.17
	T9	1.37	-1.71	-	4.45	0.38	1
	T12	-2.93	-6.11	-	0.26	0.072	1
EQ5D (MCID=0.04) [20]	-T3	-0.020	-0.048	-	0.0070	0.15	1
	T0	-0.086	-0.11	-	-0.059	<0.0001	<0.0001
	T3	-0.042	-0.071	-	-0.014	0.0039	0.16
	T6	-0.036	-0.067	-	-0.0064	0.018	0.73
	T9	-0.037	-0.069	-	-0.0052	0.023	0.95
	T12	-0.047	-0.081	-	-0.012	0.0081	0.34
BRAF-MDQ (MCID=2.6) [25]	-T3	1.41	-1.22	-	4.03	0.29	1
	T0	3.15	0.94	-	5.36	0.0053	0.68
	T3	3.25	0.60	-	5.90	0.016	1
	T6	4.33	1.85	-	6.82	0.0006	0.026
	T9	1.58	-1.32	-	4.49	0.29	1
	T12	1.76	-1.14	-	4.66	0.23	1

[a]n=42. BRAF-MDQ: Bristol Rheumatoid Arthritis fatigue multidimensional questionnaire; CI: confidence interval; DAS: disease activity score; EQ5D: European quality of life with 5 dimensions; HADS: hospital anxiety and depression scale; HAQ-DI: health assessment questionnaire disability index; MCID: minimal clinically important difference; MCS: mental component scale; PCS: physical component scale; SF36: short form 36; VAS: visual analogue scale.

Functional ability

Functional ability was 0.69(sd 0.61) at T0 in the flare group, and 0.47 (sd 0.56) in the non-flare group(figure 2B). When we visually compare the flare and non-flare group, we observed a difference that already starts six months prior to a flare and lasts until the end of the follow-up period. Not surprisingly the overall effect of a flare on the HAQ-DI was significant(p=0.0003, table 2). However, when comparing the separate time-points to the norm, a significant difference was only observed at the moment of flare and 3 months thereafter. When taking uncorrected p-values into account, the effect would last longer, namely up to 9 months. However, the difference with norm values was never above the MCID of 0.22(table 3)[19].

Health status

For the health status we compared the flare group with the non-flare group based on the physical (PCS) and mental (MCS) component score of the SF36(figure 2C, D).

Mean PCS was 36.0(sd 12.5) in the flare group and 42.8(sd 11.0) in the non-flare group. The mean MCS was respectively 55.8(sd 9.5) and 56.1(sd 10.0) in the flare and non-flare group. The overall effect of flare was not significant for the MCS(p=1), but it was for the PCS(p=0.0004). If we compare the separate time-points to norm values, a significant effect was only present at the moment of flare(table 3)[23]. Using the uncorrected p-values, there was a significant and also a clinically meaningful effect, which lasted up to 6 months after a flare (MCID SF36 PCS 3-5).

Quality of life

Quality of life shows a small dell in the graph at the moment of flare(figure 2E). The mean EQ-index at T0 was, respectively 0.75(sd 0.21) and 0.85(sd 0.13) for the patients who did and did not experienced a flare. The overall effect of a flare on EQ-5D was significant($p<0.0001$, table 2), which was also seen in the separate domains($p<0.01$), except for the domain anxiety and depression (p=0.46, data not shown). This significant effect was only seen at the moment of flare, which also exceeded the MCID threshold of 0.04[20]. If we look at the uncorrected p-values, there was a significant effect that lasted >12 months with an effect size ≥MCID for all significant time-points(table 3) [20].

Fatigue

At T0 we encountered a mean fatigue score of 19.6(sd 11.5) in the flare group and 15.7(sd 13.1) in the non-flare group(figure 2F). The effect of a flare on fatigue was significant(p=0.042, table 2). However, when comparing separate time-points the corrected p-values were not significant, while the uncorrected p-values showed a duration of 6 months. During this time period the difference with norm also exceeded the MCID of 2.6(table 3)[25].

Anxiety and depression

At visual inspection of the anxiety and depression graphs an erratic course of the scores is observed(figure 2G, H). At the moment of flare the mean anxiety scores were 3.63(sd 2.89) and 3.25(sd 2.96) for the flare and non-flare group. Mean depression scores were respectively 2.44(sd 2.22) and 2.32(sd 3.36) for the flare and non-flare group. Depression as well as anxiety scores were not influenced by a flare(p=1 for both scores, table 2).

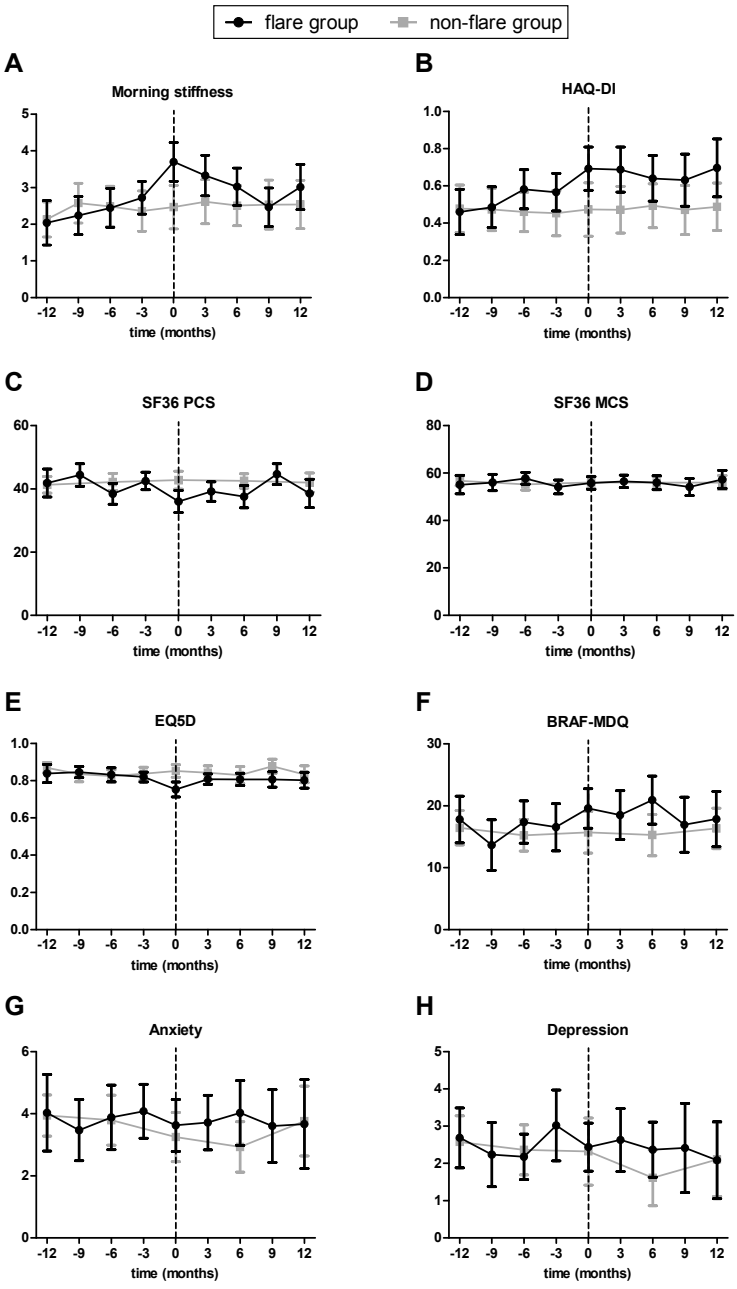


Figure 2 Patient reported outcome measures (PROMs). EQ5D: European quality of life with 5 dimensions; HAQ-DI: health assessment questionnaire disability index; SF36: short form 36; MCS: mental component scale; PCS: physical component scale.

Worker productivity

We first determined how many patients had payed work (figure 3A). At T0, 48% of the flare group and 59% of the non-flare group had payed work. Over time there were only minor differences in these numbers. Of the eligible working population respectively 27% and 18% of patients with and without a flare were unemployed at T0 (figure 3B). These percentages did not vary much over time. Sick leave was 6.2% in the flare group, and 2.6% in the non-flare group at T0, which was measured over the entire working population (figure 3C). Sick leave was not clearly affected by a flare, although we did saw a 10% drop in productivity in the 3 months after a flare, which was not significant (figure 3).

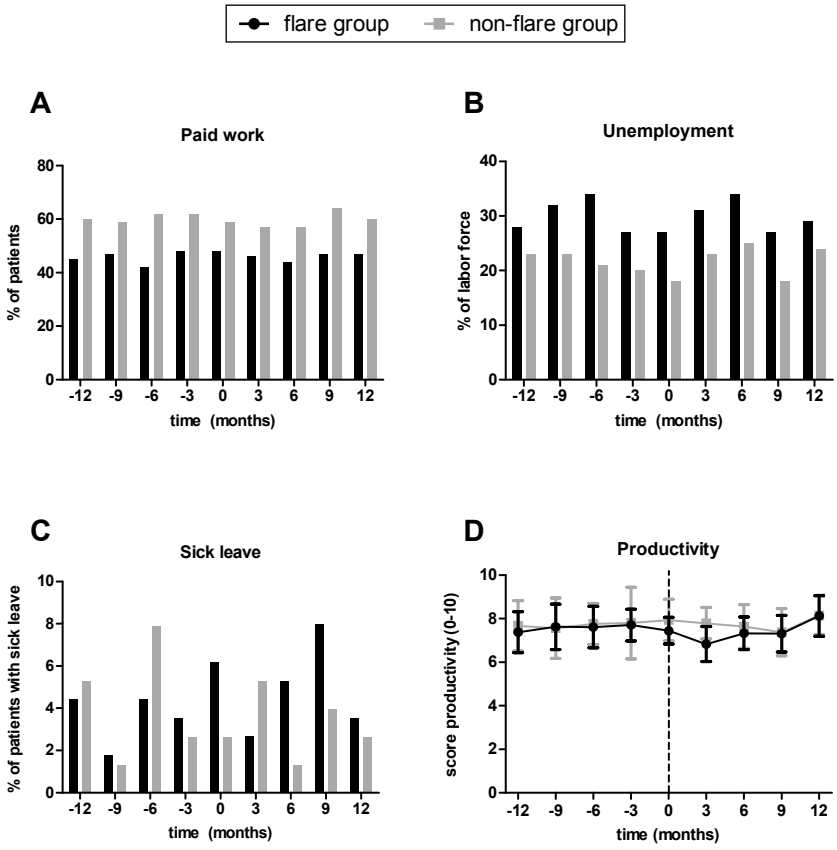


Figure 3 Worker productivity. (A) The percentage of patients with payed work, (B) unemployment as a percentage of the total labor force, (C) the amount of sick leave indicated as number of patients calling in sick within a 3 month period, (D) productivity on a scale from 0-10.

Sensitivity analysis

We performed a sensitivity analysis to evaluate the effect of different flare criteria on our PROMs(Supplementary table S1). For all flare definitions we found that DAS44, GH, morning stiffness, HAQ-DI and the EQ5D were affected(Supplementary table S1). The effect of these different flare definitions on PROMs might even be larger compared to our results.

DISCUSSION

We showed that a disease flare has a significant effect on all components of the disease activity score, but also on functional ability, quality of life and fatigue, which lasted at least 3 months. Worker productivity did not seem to be affected by a flare.

In the TARA study it was shown that tapering csDMARDs or anti-TNF in established RA patients resulted in an average flare rate of 38% during the first year of follow-up. The two tapering arms did not differ in flare rates, functional ability or quality of life [12]. Six months after the flare, 67% of the patients regained well-controlled disease [12]. These results were comparable with other tapering studies [7, 9, 11, 30]. In all these studies PROMs were merely not taken into account to assess the severity of a flare. Furthermore, it was not investigated if PROMs differed between patients with and without a flare.

However, the POET trial did show that stopping the TNF-inhibitor had a significant short-term impact on physical and mental health status compared to patients who continued their TNF-inhibitor [31]. Furthermore, the STRASS trial investigated whether the patient's perspective of a flare was the same as the physician's perspective of a flare, which was measured with the DAS28 [32]. The investigators concluded that the patient reported flare overlapped with the DAS28-based flare. The OPTIRRA trial investigators explored whether PROMs could predict a flare [33]. They showed that mental health status was independently associated with a flare during tapering. Also fatigue and functional ability were associated with a flare, but this effect disappeared after correction for possible confounders.

Although we showed a significant effect of a flare on various PROMs, this effect was not always above the MCID. For the HAQ-DI, for example, the MCID is 0.22, which was not reached in our analysis. However, the differences with the norm were statistically significant up to 6 months after a flare. Not reaching the MCID, while finding a significant differences, might be due to our assumption for the norm values. The norm was set at the average of the visits 12, 9 and 6 months prior to a flare, which was based upon the DAS44 graph. If we look at the HAQ-DI graph, we see that the HAQ-DI already worsens 6 months prior to a flare. Therefore, by taking this visit as part of the norm value, we might have underestimated the effect of flare on the HAQ-DI. For the EQ5D and the SF36 PCS we can apply a reverse reasoning of the foregoing explanation. For both PROMs we only found a significant difference at the moment of flare, while

the MCID was reached for almost every time-point after the flare, which indicates that a disease flare might have great impact on individual patients. Moreover, we corrected for multiple testing, which might have canceled out a possible meaningful effect and, therefore, underestimated the significance of our results. On the other hand, our sensitivity analysis showed similar findings for different flare definitions, which strengthens our current findings.

Strengths of the current study include the completeness of the data, including containment of recommended outcomes measures by ICHOM and OMERACT [14, 34]. Furthermore the TARA trial used a gradual tapering scheme combined with a treat-to-target approach. Therefore, we think this is an ideal trial to investigate the effect of a flare on PROMs.

Limitations of this study were that it is a post-hoc analysis. However, due to our statistical approach in which we compared patients with their own norm values, we think we can still report valid results. The results on worker productivity on the other hand are less reliable, because of the low occurrence of absenteeism and presenteeism, giving rise to a potential power issue. Furthermore, the TARA trial only had a follow-up period of 2 years, whereby potential long term effects could not be determined. For some of the investigated PROMs we already saw a long lasting effect (>6 months). Ideally, we would like to know exactly how long aforementioned effects are present, but unfortunately we do not have the data for this. There is also not always consensus about the MCIDs for specific PROMs. We used known MCIDs from the literature to place our result into perspective, but it is debatable if those values are correct.

Recently, there has been some debate on the measurement of morning stiffness, and efforts are made to create a validated PROM according to OMERACT guidelines. [35, 36]. Current used measures do not capture all aspects that are involved with morning stiffness due to RA disease activity. However, the OMERACT working group does advice not to use morning stiffness duration as outcome, because it is very aspecific.[35] Fortunately, we used the severity of morning stiffness as outcome in our analyses, but one should be cautious when interpreting these outcomes.

Due to the long-lasting effect of a flare on a patient's live, it would be ideal if we were able to predict who can safely taper medication. Current tapering strategies are based upon a trial-and-error approach, which leads to high flare rates. Our study

showed that some PROMs already worsen before a flare occurs, i.e. HAQ-DI, severity of morning stiffness and the DAS44, which might be useful for flare prevention during tapering. These changes before the actual flare occurred were all non-significant, still it indicates that patients already have more complaints before the actual flare was objectified by the treating physician. Therefore, the results of this study could be used for future research to establish a more personalized tapering approach, even though prediction of flares is not yet possible.

In conclusion, a disease flare has a significant effect on patients' lives. A disease flare affects functional ability, quality of life, fatigue, and all components of the disease activity score. The largest effect was seen in the physical outcomes, and lasted 6 months. Although on a group level the effect size for several PROMs did not exceed the specific MCID, a disease flare can still be of great importance for individual patients.

Table S1 Overall differences between norm and moments thereafter for different flare criteria.

	ΔDAS28-CRP > 0.6		**DAS28 > 2.6 & ΔDAS28 > 0.6**		**DAS28 > 3.2 & ΔDAS28 > 0.6**		**DAS28 > 3.2**	
Patient reported outcomes	P-value	Bonferroni corrected p-value*	P-value	Bonferroni corrected p-value*	P-value	Bonferroni corrected p-value*	P-value	Bonferroni corrected p-value*
DAS44	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
GH	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Morning stiffness	0.0050	0.055	0.0003	0.0032	0.0002	0.0021	0.0018	0.02
HAQ-DI	0.0043	0.047	0.019	0.21	0.0032	0.036	0.0013	0.014
EQ5D	0.0001	0.0006	<0.0001	0.0001	<0.0001	0.00015	0.0001	0.0011
BRAF-MDQ	0.13	1	0.18	1	0.042	0.46	0.24	1
HADS anxiety	1	1	0.28	1	0.35	1	0.91	1
HADS depression	0.82	1	0.86	1	0.02	0.22	0.66	1
SF36 MCS	0.62	1	0.41	1	0.37	1	0.55	1
SF36 PCS	0.12	1	0.42	1	0.40	1	0.087	0.96
Worker productivity	0.23	1	0.47	1	0.012	0.13	0.011	0.12

*n=11. BRAF-MDQ: Bristol Rheumatoid Arthritis fatigue multidimensional questionnaire; DAS: disease activity score; EQ5D: European quality of life with 5 dimensions; GH: visual analogue scale (VAS) general health; HADS: hospital anxiety and depression scale; HAQ-DI: health assessment questionnaire disability index; MCS: mental component scale; PCS: physical component scale; SF36: short form 36; VAS: visual analogue scale.

LITERATURE

1. Nieuwenhuis WP, de Wit MP, Boonen A, van der Helm-van Mil AH: Changes in the clinical presentation of patients with rheumatoid arthritis from the early 1990s to the years 2010: earlier identification but more severe patient reported outcomes. *Ann Rheum Dis* 2016, 75(11):2054-2056.
2. Haugeberg G, Boyesen P, Helgetveit K, Proven A: Clinical and Radiographic Outcomes in Patients Diagnosed with Early Rheumatoid Arthritis in the First Years of the Biologic Treatment Era: A 10-year Prospective Observational Study. *J Rheumatol* 2015, 42(12):2279-2287.
3. Lenert A, Lenert P: Tapering biologics in rheumatoid arthritis: a pragmatic approach for clinical practice. *Clin Rheumatol* 2017, 36(1):1-8.
4. Kuijper TM, Lamers-Karnebeek FB, Jacobs JW, Hazes JM, Luime JJ: Flare Rate in Patients with Rheumatoid Arthritis in Low Disease Activity or Remission When Tapering or Stopping Synthetic or Biologic DMARD: A Systematic Review. *J Rheumatol* 2015, 42(11):2012-2022.
5. Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC, Vaysbrot E, McNaughton C, Osani M, Shmerling RH *et al*: 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol* 2016, 68(1):1-26.
6. Smolen JS, Landewe R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, Nam J, Ramiro S, Voshaar M, van Vollenhoven R *et al*: EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017, 76(6):960-977.
7. Ghiti Moghadam M, Vonkeman HE, Ten Klooster PM, Tekstra J, van Schaardenburg D, Starmans-Kool M, Brouwer E, Bos R, Lems WF, Colin EM *et al*: Stopping Tumor Necrosis Factor Inhibitor Treatment in Patients With Established Rheumatoid Arthritis in Remission or With Stable Low Disease Activity: A Pragmatic Multicenter, Open-Label Randomized Controlled Trial. *Arthritis Rheumatol* 2016, 68(8):1810-1817.
8. Fautrel B, den Broeder AA: De-intensifying treatment in established rheumatoid arthritis (RA): Why, how, when and in whom can DMARDs be tapered? *Best Pract Res Clin Rheumatol* 2015, 29(4):550-565.
9. van Herwaarden N, den Broeder AA, Jacobs W, van der Maas A, Bijlsma JW, van Vollenhoven RF, van den Bemt BJ: Down-titration and discontinuation strategies of tumor necrosis factor-blocking agents for rheumatoid arthritis in patients with low disease activity. *Cochrane Database Syst Rev* 2014(9):CD010455.
10. Haschka J, Englbrecht M, Hueber AJ, Manger B, Kleyer A, Reiser M, Finzel S, Tony H-P, Kleinert S, Feuchtenberger 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* 2016, 75(1):45-51.
11. Fautrel B, Pham T, Alfaiate T, Gandjbakhch F, Foltz V, Morel J, Dernis E, Gaudin P, Brocq O, Solau-Gervais E *et al*: Step-down strategy of spacing TNF-blocker injections for established rheumatoid arthritis in remission: results of the multicentre non-inferiority randomised open-label controlled trial (STRASS: Spacing of TNF-blocker injections in Rheumatoid ArthritiS Study). *Ann Rheum Dis* 2016, 75(1):59-67.
12. van Mulligen E, de Jong PHP, Kuijper TM, van der Ven M, Appels C, Bijkerk C, Harbers JB, de Man Y, Molenaar THE, Tchetverikov I *et al*: Gradual tapering TNF inhibitors versus conventional synthetic DMARDs after achieving controlled disease in patients with rheumatoid arthritis: first-year results of the randomised controlled TARA study. *Ann Rheum Dis* 2019, 78(6):746-753.
13. Heller JE, Shadick NA: Outcomes in rheuma-

toid arthritis: incorporating the patient perspective. *Curr Opin Rheumatol* 2007, 19(2):101-105.

14. Oude Voshaar MAH, Das Gupta Z, Bijlsma JWJ, Boonen A, Chau J, Courvoisier DS, Curtis JR, Ellis B, Ernestam S, Gossec L *et al*: The International Consortium for Health Outcome Measurement (ICHOM) Set of Outcomes that Matter to People Living with Inflammatory Arthritis Consensus from an international Working Group. *Arthritis Care Res (Hoboken)* 2018.
15. van Mulligen E, de Jong PHP, Kuijper TM, van der Ven M, Appels C, Bijkerk C, Harbers JB, de Man Y, Molenaar THE, Tchetverikov I *et al*: Gradual tapering TNF inhibitors versus conventional synthetic DMARDs after achieving controlled disease in patients with rheumatoid arthritis: first-year results of the randomised controlled TARA study. *Ann Rheum Dis* 2019.
16. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, van Rijswijk MH, van de Putte LB: Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990, 49(11):916-920.
17. Anderson JK, Zimmerman L, Caplan L, Michaud K: Measures of rheumatoid arthritis disease activity: Patient (PtGA) and Provider (PrGA) Global Assessment of Disease Activity, Disease Activity Score (DAS) and Disease Activity Score with 28-Joint Counts (DAS28), Simplified Disease Activity Index (SDAI), Clinical Disease Activity Index (CDAI), Patient Activity Score (PAS) and Patient Activity Score-II (PASII), Routine Assessment of Patient Index Data (RAPID), Rheumatoid Arthritis Disease Activity Index (RADAI) and Rheumatoid Arthritis Disease Activity Index-5 (RADAI-5), Chronic Arthritis Systemic Index (CASI), Patient-Based Disease Activity Score With ESR (PDAS1) and Patient-Based Disease Activity Score without ESR (PDAS2), and Mean Overall Index for Rheumatoid Arthritis (MOI-RA). *Arthritis Care Res (Hoboken)* 2011, 63 Suppl 11:S14-36.
18. Siegert CE, Vleming LJ, Vandenbroucke JP, Cats A: Measurement of disability in Dutch rheumatoid arthritis patients. *Clin Rheumatol* 1984, 3(3):305-309.
19. Redelmeier DA, Lorig K: Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology. *Arch Intern Med* 1993, 153(11):1337-1342.
20. Luo N, Johnson J, Coons SJ: Using instrument-defined health state transitions to estimate minimally important differences for four preference-based health-related quality of life instruments. *Med Care* 2010, 48(4):365-371.
21. EuroQol G: EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy* 1990, 16(3):199-208.
22. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A: Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *Br J Rheumatol* 1997, 36(5):551-559.
23. Ware JE, Jr., Sherbourne CD: The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992, 30(6):473-483.
24. Nicklin J, Cramp F, Kirwan J, Urban M, Hewlett S: Collaboration with patients in the design of patient-reported outcome measures: capturing the experience of fatigue in rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2010, 62(11):1552-1558.
25. Dures EK, Hewlett SE, Cramp FA, Greenwood R, Nicklin JK, Urban M, Kirwan JR: Reliability and sensitivity to change of the Bristol Rheumatoid Arthritis Fatigue scales. *Rheumatology (Oxford)* 2013, 52(10):1832-1839.

26. Zigmond AS, Snaith RP: The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983, 67(6):361-370.
27. Puhan MA, Frey M, Buchi S, Schunemann HJ: The minimal important difference of the hospital anxiety and depression scale in patients with chronic obstructive pulmonary disease. *Health Qual Life Outcomes* 2008, 6:46.
28. Lemay KR, Tulloch HE, Pipe AL, Reed JL: Establishing the Minimal Clinically Important Difference for the Hospital Anxiety and Depression Scale in Patients With Cardiovascular Disease. *J Cardiopulm Rehabil Prev* 2018.
29. Bouwmans C, Krol M, Severens H, Koopmanschap M, Brouwer W, Hakkaart-van Roijen L: The iMTA Productivity Cost Questionnaire: A Standardized Instrument for Measuring and Valuing Health-Related Productivity Losses. *Value Health* 2015, 18(6):753-758.
30. Smolen JS, Nash P, Durez P, Hall S, Ilivanova E, Irazoque-Palazuelos F, Miranda P, Park MC, Pavelka K, Pedersen R *et al*: Maintenance, reduction, or withdrawal of etanercept after treatment with etanercept and methotrexate in patients with moderate rheumatoid arthritis (PRESERVE): a randomised controlled trial. *Lancet* 2013, 381(9870):918-929.
31. Ghiti Moghadam M, Ten Klooster PM, Vonkeman HE, Kneepkens EL, Klaasen R, Stolk JN, Tchetverikov I, Vreugdenhil SA, van Woerkom JM, Goekoop-Ruiterman YPM *et al*: Impact of Stopping Tumor Necrosis Factor Inhibitors on Rheumatoid Arthritis Patients' Burden of Disease. *Arthritis Care Res (Hoboken)* 2018, 70(4):516-524.
32. Portier A, Gossec L, Tubach F, Alfaiate T, Pham T, Saraux A, Soubrier M, Schaeverbeke T, Maillefert JF, Mariette X *et al*: Patient-perceived flares in rheumatoid arthritis: A sub-analysis of the STRASS treatment tapering strategy trial. *Joint Bone Spine* 2017, 84(5):577-581.
33. Bechman K, Sin FE, Ibrahim F, Norton S, Matcham F, Scott DL, Cope A, Galloway J: Mental health, fatigue and function are associated with increased risk of disease flare following TNF inhibitor tapering in patients with rheumatoid arthritis: an exploratory analysis of data from the Optimizing TNF Tapering in RA (OPTTIRA) trial. *RMD Open* 2018, 4(1):e000676.
34. van Tuyl LH, Sadlonova M, Davis B, Flurey C, Goel N, Hewlett SE, Hill CL, Hoogland W, Kirwan JR, van Schaardenburg D *et al*: Remission in Rheumatoid Arthritis: Working Toward Incorporation of the Patient Perspective at OMERACT 12. *J Rheumatol* 2016, 43(1):203-207.
35. Craig ET, Orbai AM, Mackie S, Bartlett SJ, Bingham CO, 3rd, Goodman S, Hill C, Holt R, Leong A, Karyekar C *et al*: Advancing Stiffness Measurement in Rheumatic Disease: Report from the Stiffness Special Interest Group at OMERACT 2018. *J Rheumatol* 2019, 46(10):1374-1378.
36. Halls S, Sinnathurai P, Hewlett S, Mackie SL, March L, Bartlett SJ, Bingham CO, 3rd, Alten R, Campbell I, Hill CL *et al*: Stiffness Is the Cardinal Symptom of Inflammatory Musculoskeletal Diseases, Yet Still Variably Measured: Report from the OMERACT 2016 Stiffness Special Interest Group. *J Rheumatol* 2017, 44(12):1904-1910.

CHAPTER 9

TWO-YEAR COST-EFFECTIVENESS BETWEEN TWO GRADUAL TAPERING STRATEGIES IN RHEUMATOID ARTHRITIS: COST-UTILITY ANALYSIS OF THE TARA TRIAL

E van Mulligen, AEAM Weel, TM Kuijper, NHAM Denissen, AH Gerards,

M de Jager, WK Lam, JMW Hazes, AHM van der Helm – van Mil, PHP de Jong*,

JJ Luime*

Ann Rheum Dis 2020;79(12):1550-1556.

ABSTRACT

Objective The aim of the current study was to evaluate the two year cost-utility ratio between tapering csDMARD first followed by the TNF-inhibitor, or vice versa, in rheumatoid arthritis patients.

Methods Two-year data of the TARA tapering trial were used. RA patients who used both a csDMARD and a TNF-inhibitor and had a well-controlled disease (DAS≤2.4&SJC≤1) for at least 3 months, were randomized into gradual tapering the csDMARD first followed by the TNF-inhibitor, or vice versa. Quality-adjusted life years (QALYs) were derived from the EQ5D. Health care and productivity costs were calculated with data from patient records and questionnaires. The incremental cost-effectiveness ratio (ICER) and the incremental net monetary benefit (iNMB) were used to assess cost-effectiveness between both tapering strategies.

Results 94 patients started tapering their TNF-inhibitor first, while the other 95 tapered their csDMARD first. QALYs (sd) were, respectively, 1.64 (0.22) and 1.65 (0.22). Medication costs were significantly lower in the patients who tapered the TNF-inhibitor first, while indirect cost were higher due to more productivity loss (p=0.10). Therefore, total costs (sd) were €38,833 (€39,616) for tapering csDMARDs first, and €39,442 (€47,271) for tapering the TNF-inhibitor (p=0.88). For willingness-to-pay (WTP) levels <€83,800 tapering the csDMARD first has the highest probability of being cost-effective, while for WTP levels >€83,800 tapering the TNF-inhibitor first has the highest probability.

Conclusion Our economic evaluation shows that costs are similar for both tapering strategies. Regardless of the WTP, tapering either the TNF-inhibitor or the csDMARD first is equally cost-effective.

INTRODUCTION

The optimal management for Rheumatoid Arthritis (RA) comprises an early, intensive and treat-to-target management approach, which has the highest chance of inducing remission and preventing joint damage.[1, 2] In case of sustained remission, tapering of treatment can be considered to reduce side-effects and save costs.[3] In the Netherlands more than 300 million euros are spent on the use of biological therapy for rheumatic diseases.[4] On the other hand, treatment with conventional synthetic disease-modifying anti-rheumatic drugs (csDMARDs) generally costs only one tenth of the cost of a biological.[5] Although rheumatologists carefully consider initiation of biologicals, uniform tapering decisions are lacking, and therefore biological tapering is not always directly performed when sustained remission is achieved.[6] Tapering of biologicals could reduce health care costs.

In the TARA trial two tapering strategies were compared, namely tapering the TNF-inhibitor first followed by the csDMARD, or vice versa. Within the first year, in which either the TNF-inhibitor or the csDMARD was gradually tapered within 6 months, there were no significant differences in flare rates, disease activity, functional ability and quality of life, although we did observe numerical differences (10% in flare rates), and less patients in clinical remission.[7] From a clinical viewpoint one could argue that the order of tapering is not relevant. On the other hand, TNF-inhibitors are far more expensive than csDMARDs, therefore from a health economics perspective it is more sensible to taper the TNF-inhibitor first. Previous studies already showed that tapering biologicals leads to a reduction of medication and medical consumption costs, also known as direct costs, but could also result in a decrease in quality of life.[8-11] Tapering of medication might lead to an increase in disease activity and consequently to a disease flare. This could lead to more pain and disability, possibly resulting in more productivity loss and sick leave. However, not much is known about aforementioned possible effects. Furthermore, none of the aforementioned studies compared two active tapering strategies.[11, 12]

Moreover, a previous study already showed that disease flares have a significant impact on patients' lives, which among other things could lead to productivity loss.[13] As mentioned earlier, the effect of a flare on societal costs is not known. Nor do we know whether the health care (direct) cost reduction due to tapering treatment outweigh the possible increase in productivity (indirect) costs.

Therefore, our aim is to investigate which gradual tapering strategy has the best cost-utility ratio over a period of two years. Furthermore, we want to explore the effect of tapering on both medical and societal costs.

PATIENTS AND METHODS

Patients

For this study data were used from the Tapering strategies in Rheumatoid Arthritis (TARA) trial (NTR2754). TARA, a multicenter, single-blinded trial was carried out in twelve rheumatology centers in the Netherlands between September 2011 and July 2016. Medical ethics committees at each participating center approved the study protocol and all patients gave written informed consent before inclusion.

Primary aims of the TARA study were to assess effectiveness and cost-effectiveness of two tapering strategies, from a societal perspective. An extended description of the TARA study and clinical effectiveness outcomes can be found elsewhere.[7, 14] Inclusion criteria for the TARA trial were: adult RA patients, with a well-controlled disease, defined as a disease activity score (DAS44)≤2.4 and a swollen joint count (SJC)≤1 at two consecutive time points within a 3-month interval, who were using a combination of a csDMARD and a TNF-inhibitor.

Randomization and masking

Patients were randomized using minimization randomization stratified for center into tapering the csDMARD in the first year followed by tapering the TNF-inhibitor in the second year, or vice versa. No other factors were used for the minimization randomization. Trained research nurses, blinded to the allocated treatment arm throughout the study, examined the patients.

Design

The csDMARD and TNF-inhibitor were both gradually tapered in three steps. csDMARD tapering was realised by cutting the dosage into half, a quarter and thereafter it was stopped. The TNF-inhibitor was tapered by doubling the interval between gifts, followed by cutting the dosage into half, and thereafter it was stopped. If patients remained flare-free, the first drug was completely tapered after 6 months.

Both tapering strategies had a treat-to-target approach with three-monthly visits. At each visit patients were assessed whether they maintained low disease activity (DAS≤2.4) while tapering their medication. If a disease flare occurred, defined as a DAS>2.4 and/or SJC>1, tapering was stopped and the last effective treatment, when the patient still had well-controlled disease, was restarted. No further attempts were taken to taper medication. Treatment was intensified at each visit until low disease activity was reached again.

Concurrent treatment with non-steroidal anti-inflammatory drugs and intra-articular glucocorticoid injections were allowed. In case of a flare, one intra-muscular glucocorticoid injection was allowed to be given as bridging therapy, in addition to switching to the last effective dosage of the csDMARD or TNF-inhibitor.

Effectiveness and cost assessment

The primary outcome of the TARA study was the number of disease flares. For the cost-effectiveness the main outcome was the incremental cost-effectiveness ratio (ICER). The ICER is the ratio of the difference in costs compared to the difference in quality adjusted life years (QALYs) between both tapering strategies. Costs per QALY were calculated, since coverage of prescribed drugs by Dutch health insurance companies depends on this outcome. The required threshold per additional QALY gained to be funded for a new intervention in the Netherlands is €50,000.[15-17] QALYs express the impact of the disease on patients' health over time. Living in perfect health for one year corresponds to 1 QALY, living in perfect health for two years corresponds to 2 QALYs. Zero QALYs reflects death at baseline.[18] QALYs were determined by calculating the area under the curve of the EuroQol questionnaire with 5 dimensions (EQ-5D) with 3 levels over a two year period.[19]

Total costs are divided into health care (direct) and productivity (indirect) costs. We analyzed health care and productivity costs from a societal perspective. Health care costs are the costs of treatment and medical consumption, whereas productivity costs are costs due to presenteeism, i.e. working while sick, and absenteeism, i.e. sick leave and unemployment.[20]

Medication costs were calculated from doses reported in the patients' case records, valued according to the Dutch college of health insurances (supplementary table S1).[5] Duration of hospitalizations and admission diagnosis were recorded every three months with the iMTA medical consumption questionnaire. Medical consumption,

including hospital admissions, was valued at Dutch standard prices, except for costs of complementary and alternative medicine, which were based upon American data, because no Dutch data are available (supplementary table S2).[21, 22]

Productivity costs included absenteeism, such as sick leave and reduction in work time, and presenteeism, including working while sick. Every three months patients filled out the iMTA productivity cost questionnaires (iPCQ).[23] The friction cost method was used to calculate the productivity costs, which assumes replaceability of every employee in time.[19] The friction cost period is the time between the start of long-term sick leave, and filling the position again. Costs due to sick leave are solely counted during this period, which encompasses 85 days in the Netherlands.[24] Productivity losses were valued at age- and sex-dependent standard hourly costs (supplementary table S3).[25, 26] All prices were obtained for the year 2019. Costs were not discounted, because of only two years of follow-up.

Willingness-to-pay

To help decide which tapering strategy has the highest chance of being cost-effective, two indicators were used. First, the cost-effectiveness acceptability curves were derived to show the probability of each tapering strategy being cost-effective at different levels of willingness-to-pay (WTP) thresholds in comparison with each other.[27] Second, the incremental net monetary benefit (iNMB) was used to express the incremental value of the tapering strategies in monetary terms at different levels of willingness-to-pay per QALY. This results in an alternate measure which reports on cost-effectiveness without using the ICER. The iNMB was calculated as the incremental benefit times different levels of WTP, minus the incremental costs. A positive iNMB indicates that the tapering the TNF-inhibitor first is cost-effective compared to tapering the csDMARD first.[27]

Statistical analysis

The cost-effectiveness analysis follows a superiority design. Sample size calculation was based on the number of disease flares after one year, which was described previously.[7] All analysis were performed following an intention-to-treat approach.

After two years of follow-up, 13/94 (13.8%) in the tapering the csDMARD first group had dropped out, versus 9/95 (9.5%) in the tapering the TNF-inhibitor first group. Furthermore 7.6% of patients who completed the trial did not completely fill out the questionnaires. Multiple imputations with chained equations (MICE), with 40

imputations, were used to handle missing data in baseline variables as well as in the follow-up data.[28] An imputation regression model was constructed to impute EQ-5D, unemployment, loss of productivity due to sick leave (absenteeism) and not fully functioning (presenteeism) and the (decrease in) number of working hours.

For EQ-5D, presenteeism and the amount of working hours linear regression was used. The percentages of missingness for these variables were, respectively, 14.9%, 6.7%, 18.6%. For presenteeism we log transformed the variable and used linear regression to impute values. For unemployment (13.6% missing values) we used logistic regression, and for sick leave (7.9% missing values) we used a Poisson regression model. The choice of imputation models were based on the distribution of the individual variables. In the regression models we used age, gender, baseline values, and the tapering strategy as independent variables. Differences between imputed data, created with aforementioned models, and complete cases were minimal and showed that our imputation models are reliable (supplemental table S4).

The main outcome was the incremental cost-effectiveness ratio (ICER). A probabilistic sensitivity analysis for the estimation of the ICER was performed by bootstrapping with 1000 iterations using a Monte Carlo simulation. Results were plotted in a cost-effectiveness plane and were used to estimate the 95% confidence interval (CI) of the ICER.

Differences in outcomes between groups were analyzed with linear regression models, and to account for stratified randomization by center, intercepts for each center were included.

All data was analyzed using STATA 15, using a value of $p \leq 0.05$ as the level of statistical significance.

RESULTS

Patients

A total of 189 patients were randomly assigned to taper the csDMARD (n=94) or TNF-inhibitor (n=95) first. Over two years, 22 patients (11.6%) withdrew from the study, resulting in 167 patients with a complete follow-up. At baseline, patients had an average symptom duration of 6.8 years and were predominantly female (66.1%) with an average age of 56.6 years (table 1). The majority of patients (55%) used etanercept as their TNF-inhibitor. At baseline, 47 (25%) of patients were aged above 65, which

was the average age of retirement in the Netherlands in 2018.[29] Of the 142 patients under 65, 99 patients (70%) had paid work at baseline (table 1).

Table 1 Baseline characteristics of both tapering groups.

	Tapering csDMARDs first (n=94)	Tapering TNF-inhibitor first (n=95)
Demographic		
Age (years), mean (sd)	55.9 (14)	57.2 (11)
Aged above 65, n (%)	22 (23)	25 (26)
Gender, female, n (%)	67 (71)	58 (61)
Quality of life		
EQ5D index, mean (sd)	0.86 (0.12)	0.87 (0.11)
Disease characteristics		
Symptom duration (years), median (IQR)	6.0 (4.3-8.5)	6.3 (4.1-8.9)
RF positive, n (%)	49 (57)	56 (64)
ACPA positive, n (%)	61 (72)	65 (75)
DAS, mean (sd)	1.1 (0.6)	1.0 (0.5)
Use of csDMARDs [a]		
MTX, n (%)	89 (95)	84 (88)
SASP, n (%)	10 (11)	12 (13)
HCQ, n (%)	24 (26)	37 (39)
Leflunomide, n (%)	2 (2)	4 (4)
Use of TNF-inhibitors		
Etanercept, n (%)	52 (55)	52 (55)
Adalimumab, n (%)	36 (39)	40 (43)
Other, n (%) [b]	6 (7)	3 (3)
Worker related outcomes		
Paid work, n (%) [c]	47 (61)	52 (68)
Working hours per week, mean (sd)	28 (8)	29 (11)

[a] some patients used a combination of csDMARDs, [b] certolizumab or golimumab, [c] number of patients with paid work and aged under 65. ACPA: anti-citrullinated protein antibody; csDMARDs: conventional synthetic disease modifying anti-rheumatic drugs; DAS: disease activity score based on 44 joints; EQ5D: European Quality of life questionnaire with 5 dimensions; HCQ: hydroxychloroquine; IQR: interquartile range; MTX: methotrexate; RF: rheumatoid factor; SASP: salazopyrine; sd: standard deviation.

Health care costs

Mean health care costs (sd) were €22,484 (€8,069) for tapering the csDMARD first and €13,616 (€9,162) for tapering the TNF-inhibitor first ($p<0.001$)(table 2). Respectively, 86% and 71% of health care costs were medication costs. The faster savings due to less TNF-inhibitor use within the group that tapered the TNF-inhibitor first was the main driver of the difference in direct costs. Within the group who tapered the

csDMARDs first, 81 (86%) were using full dose TNF-inhibitor after 12 months, and 32 (34%) patients after 24 months. In the TNF-inhibitor tapering first group this was 16 (17%) after 12 months, and 25 (26%) after 24 months.

Table 2 Health care costs over two years of follow-up in the TARA study according to intention-to-treat.

	Tapering csDMARDs first (n=94)		**Tapering TNF-inhibitor first (n=95)**	
	Number of visits, mean (sd)	**Mean costs (sd)**	**Number of visits, mean (sd)**	**Mean costs (sd)**
Medication *				
csDMARDs *		€436 (€87)		€972 (€123)
TNF-inhibitor *		€19,417 (€738)		€9,673 (€863)
Prednisone		€2.46 (€0.54)		€2.84 (€0.59)
Medical consumption				
Hospitalization	13 [a]	€326 (€1313)	15 [a]	€558 (€2271)
Standard health care				
Primary care physician	7.7 (9)	€260 (€302)	8.9 (9)	€303 (€318)
Specialist	12.0 (6)	€1,153 (€647)	12 (6)	€1,203 (€738)
Psychologist	0.5 (2)	€18 (€83)	1.2 (8)	€40 (€266)
Paramedical care				
Physical therapy	14.4 (32)	€506 (€1,110)	15.9 (31)	€554 (€1,063)
Dietitian	0.46 (2)	€14 (€62)	0.040 (0.3)	€1.31 (€8.95)
Social worker	0.14 (0.6)	€9.40 (€41)	0.20 (0.8)	€14 (€52)
Speech therapist	0.04 (0.3)	€1.32 (€10)	0.02 (0.2)	€0.65 (€6.36)
Alternative medicine				
Homeopathy	0.83 (3)	€26 (€97)	0.44 (2)	€14 (€67)
Total health care costs, mean (sd)		€22,484 (€8,069)		€13,616 (€9,162)

* $p<0.001$ (linear regression adjusted for stratified randomisation. [a] Number reflects the number of patients who got hospitalized within the two years of follow-up. csDMARDs: conventional synthetic disease modifying anti-rheumatic drugs; sd: standard deviation.

Productivity costs

Average productivity costs (sd) for tapering csDMARDs first and TNF-inhibitor first were, respectively €16,349 (€38,277) and €25,826 (€46,289)(p=0.10)(table 3, 4). Within the two years of follow-up 20 (43%) patients with paid work called in sick with an average duration of 9 days in the initial csDMARD tapering group versus 26 patients (50%) with an average duration of 12 days within the initial TNF-inhibitor tapering group. Of those patients, respectively 2 and 1 had long-term sickness (>3 months). Two patients

who tapered the csDMARD first became unemployed, versus six in the group who tapered the TNF-inhibitor first. The working population had an average workweek of 32 hours after 24 months of follow-up. A decrease in working hours was seen in 8 and 11 patients in respectively the csDMARD and TNF-inhibitor tapering first group. Their average workweek decreased with 15 hours in the csDMARD tapering first group and 19 hours in the TNF-inhibitor tapering first group. Within the working population 34 patients in the csDMARD tapering first group and 41 patients in the TNF-inhibitor tapering first group indicated that they had days on which they were less productive. On average, this were 5 and 6 days per month, with a mean productivity loss on these days of 28% and 26%, respectively (table 3). Sub analyses of males and females did not result in differences in productivity costs (data not shown).

Table 3 Productivity costs over two years of follow-up.

	Tapering csDMARDs first (n=94)	**Tapering TNF-inhibitor first (n=95)**
Absenteeism		
Unemployment		
Became unemployed, n (%)	2 (4)	6 (11)
Sick leave (during 2-year follow-up)		
Occurrence, n (%)	20 (21)	26 (27)
Long-term sickness, n (%)	2 (2)	1 (1)
Days absent, mean (sd) [a]	9.0 (23)	12.3 (22)
Contract hours [b]		
Working hours per week after 2 years, mean (sd)	32 (8.9)	33 (12)
Reduction of working hours per week, n (%)	8 (8)	11 (11)
Amount of reduction, hours, mean (sd) [c]	15 (11)	19 (17)
Presenteeism		
Number of patients, n (%)	34 (36)	41 (43)
- Number of days per month, mean (sd) [d]	5.3 (0.9)	6.1 (1.1)
- Average productivity loss, proportion (sd) [e]	27.9% (13%)	26.4% (15%)

[a] Only indicated when patients reported sick leave
[b] Only indicated when patients had paid work
[c] Only indicated for those with a reduction in working hours
[d] Average productivity score was only obtained for patients indicating that they had loss of productivity.
[e] Productivity loss was indicated only for the days with productivity loss for those who reported to suffer from loss of productivity.
csDMARDs: conventional synthetic disease modifying anti-rheumatic drugs, sd: standard deviation

Cost-effectiveness analysis

The mean EQ5D index (sd) after 24 months of follow-up was 0.81 (0.13) for tapering the csDMARD first, and 0.83 (0.16) for tapering the TNF-inhibitor first. Average QALYs

(sd) over two years for tapering csDMARDs first or TNF-inhibitor first were, respectively, 1.64 (0.22) and 1.65 (0.22)(table 4). Total costs (sd) were €38,833 (€39,616) for tapering csDMARDs first, and €39,442 (€47,271) for tapering the TNF-inhibitor (p=0.88)(table 4).

The ICER (95% CI) between tapering csDMARDs first minus the TNF-inhibitor first was €60,919 per QALY (-€90,638 per QALY to €212,475 per QALY), indicating that tapering TNF-inhibitor first was on average €60,919 less expensive per QALY compared to tapering the csDMARD first. However, the confidence interval is very wide due to a minimal difference in QALYs and costs between the two tapering strategies. To illustrate this the analysis of uncertainty in the estimation of the ICER was visualized with the cost-effectiveness planes for the two tapering strategies compared to each other (figure 1A). The iNMB was €1134 (95% CI €761 to €1507) in favor of tapering TNF-inhibitor first for a willingness-to-pay (WTP) level of €50,000, which is the current level of WTP in the Netherlands for treatment of RA (supplemental figure S1).[15-17] Our cost-effectiveness acceptability curve (CEAC) shows similar results (figure 1B). For WTP levels <€83,800 tapering the csDMARD first has the highest probability of being cost-effective, while for WTP levels >€83,800 tapering the TNF-inhibitor first has the highest probability**.** In between WTP levels of €53,800 and €83,800 both strategies were evenly cost-effective (probability 50%). This indicates that depending on the WTP threshold either tapering the TNF-inhibitor or csDMARD first is more cost-effective. Moreover, the CEAC shows that both lines are almost horizontal after the crossing and that the difference is small, which is due to the small differences in QALYs and costs.

Table 4 Total costs and QALYs over the 2 year follow-up period

	Tapering csDMARD first	**Tapering TNF-inhibitor first**
Total costs	€38,833 (€39,616)	€39,442 (€47,271)
Total health care costs *	€22,484 (€8,069)	€13,616 (€9,162)
• Medication *	€19,858 (€7,343)	€10,648 (€8,642)
• Medical consumption	€2,297 (€1,684)	€2,393 (€1,775)
• Hospitalization	€330 (€1,319)	€575 (€2,305)
Total productivity costs	€16,349 (€38,277)	€25,826 (€46,289)
• Absenteeism	€17,581 (€39,576)	€23,577 (€45,382)
• Presenteeism	€3,290 (€9,952)	€4,777 (€14,620)
QALYs (EQ5D, AUC), mean (sd)	1.64 (0.22)	1.65 (0.22)

* $p<0.0001$ (linear regression adjusted for stratified randomisation). All values are indicated as mean (sd). AUC: area under the curve; csDMARDs: conventional synthetic DMARDs; EQ-5D: Dutch European quality of life; QALY: quality adjusted life year; sd: standard deviation.

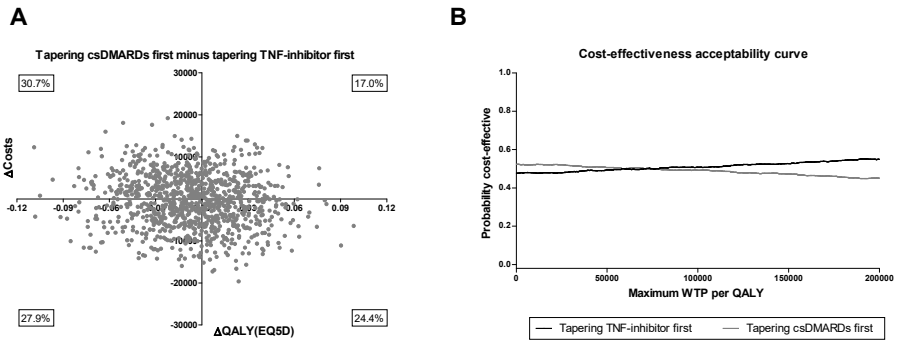


Figure 1 Summary of economic evaluation of tapering csDMARDs first minus tapering TNF-inhibitor first. (A) Results of 1000 bootstrapped replications, presented in a cost-effectiveness plane which represents uncertainty of the cost-effectiveness ratio. (B) Cost effectiveness acceptability curve for tapering csDMARDs first versus tapering TNF-inhibitor first. Results of 1000 bootstrapped replication, presented for several levels of willingness to pay, indicated per quality adjusted life year (QALY) csDMARDs: conventional synthetic DMARDs; QALY: quality adjusted life year; WTP: willingness to pay.

DISCUSSION

In this study, we showed that health care costs were significantly lower in patients who tapered the TNF-inhibitor first, but productivity costs in this group were higher due to more absenteeism and presenteeism compared to the patients who tapered the csDMARD first. The ICER (95% CI) between tapering csDMARDs first minus the TNF-inhibitor first was €60,919 per QALY (-€90,638 to €212,475). Total costs (sd) were €38,833 (€39,616) for tapering csDMARDs first, and €39,442 (€47,271) for tapering the TNF-inhibitor first (p=0.88). Depending on the WTP threshold either tapering the TNF-inhibitor or csDMARD first has the highest probability of being cost-effective.

Previous studies showed that savings on health care and societal costs could be obtained by treating to target within newly diagnosed RA patients.[30] More savings could be obtained by tapering quickly, and possibly stopping the medication when RA patients reach sustained remission. Currently, several trials have reported on the feasibilty of tapering, however cost-effectiveness analyses are scarce. A systematic review on tapering and stopping treatment in RA patients reported that only 2 out of 14 included studies performed a cost-effectiveness analysis, although costs are nowadays an important reason why tapering or stopping treatment is considered by treating rheumatologists.[31]

Previous studies reported on the cost-effectiveness of tapering or stopping medication versus a continuation group. The DRESS study for example showed a

significant cost-saving after tapering of adalimumab or etanercept, without a clinically meaningful loss in QALYs.[12] The STRASS trial also reported on cost-effectiveness. Within this trial the interval between TNF-inhibitor injections was extended and compared to a control group that continued their medication. Health care costs were significantly lower in the tapering group, but this was accompanied with a signficant loss in QALYs.[11] Although both studies also reported on productivity costs, they did not take presenteeism into account. In our study the QALYs did not differ between both tapering strategies and were comparable to the QALYs of the control groups in previous mentioned trials (DRESS 1.67 and STRASS 1.68).[11, 12]

The strengths of the current study include the randomized design. Although originally the TARA trial was powered to find a 20% difference in disease flares, cost-effectiveness was a parallel primary outcome. Also, validated outcome measures were used for the QALY calculation. Furthermore, we used real data to calculate health care and productivity costs, instead of using a model. Moreover, for calculating productivity costs we included absenteeism as well as presenteeism, thereby taking into account all costs due to productivity loss. Finally, the TARA trial is the first randomized controlled trial reporting on the cost-utilty between two gradual tapering strategies.

Some limitations should also be noted. First of all, the targeted sample size was not reached. This was due to difficulties with inclusion, and the start of another trial using the same pool of eligible patients. For the primary outcome (disease flares) we performed a sensitivity analysis, which showed similar outcomes.[7] Furthermore, the follow-up duration was only 2 years. Ideally, longterm effects of tapering and stopping treatment should be taken into account as well. In our current design, patients completely tapered their medication after 18 months, if no flare occurred. This means that we only have 6 months of follow-up when patients are in DMARD-free remission. Late flares were, therefore, not considered in our study and might change current outcomes by an increase in health care costs on the long term, but might also influence productivity costs and quality of life.

Generalizability of the current study might be difficult, since every country has its own social security and healthcare system. Also treatment prices differ. Costs of labor vary between countries, and more importantly, rules and regulations for social security regulation differ across countries. The possibility to stay at home when not feeling well is very different across countries within Europe.[32] In the Netherlands, people can call in sick without consulting a doctor, while this is obligatory in some other countries. This could cause a shift between presenteeism and absenteeism when comparing the Netherlands to other countries. Fortunately, in our current analysis

we do take into account both. Since we found that the group that tapered the TNF-inhibitor first encountered more costs due to both presenteeism and absenteeism (table 3, p=0.39 and p=0.20, respectively), we believe that our indirect costs are generalizable to all countries.

In the current study we found a significant difference in medication costs between both tapering groups. The difference in health care costs could change due to price variations of csDMARD and especially biologicals between countries. To investigate this, we performed a sensitivity analysis with varying levels of biological prices of 30%, 50%, and 200% of the prices we currently used (supplemental figure S2). Lowering biological prices was in favor of tapering the csDMARD first, while higher biological prices showed the opposite. However, biological costs are consistantly higher than csDMARD costs in any country, meaning that the direction of the medication cost difference could be generalizable to other countries. For the current analysis we used 2019 prices to make our results as relevant as possible, since the prices for biologicals have decreased dramatically.

In conclusion, medication costs are lower when the TNF-inhibitor is tapered first, but this is counterbalanced with a higher loss of productivity and, therefore, cost savings are similar for both tapering strategies. Regardless of the WTP threshold, tapering the TNF-inhibitor or csDMARD first is equally cost-effective.

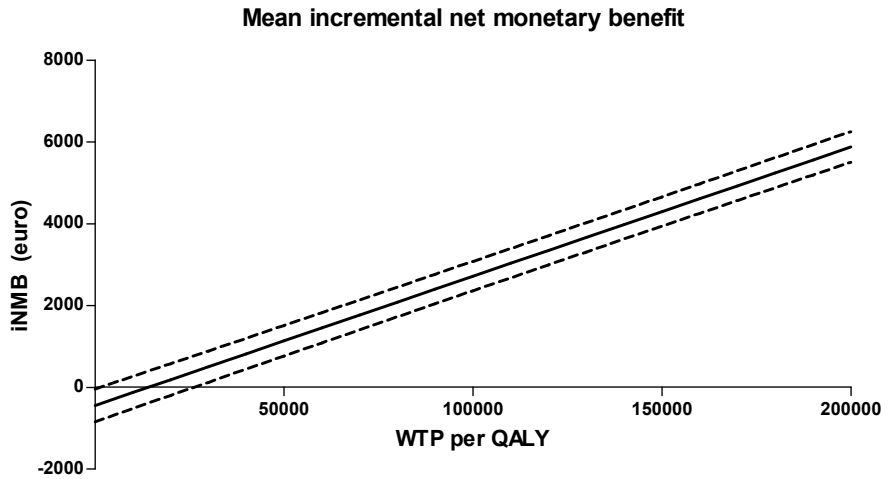


Figure S1 Mean incremental net monetary benefit (iNMB) for tapering csDMARDs minus tapering TNF-inhibitors first. Results are plotted against different levels of willingness to pay (WTP) per quality adjusted life year (QALY), and with 95% confidence intervals. The iNMB was calculated as the incremental benefit times different levels of WTP, minus the incremental costs. csDMARDs: conventional synthetic DMARDs; iNMB: incremental net monetary benefit; QALY: quality adjusted life year; WTP: willingness to pay.

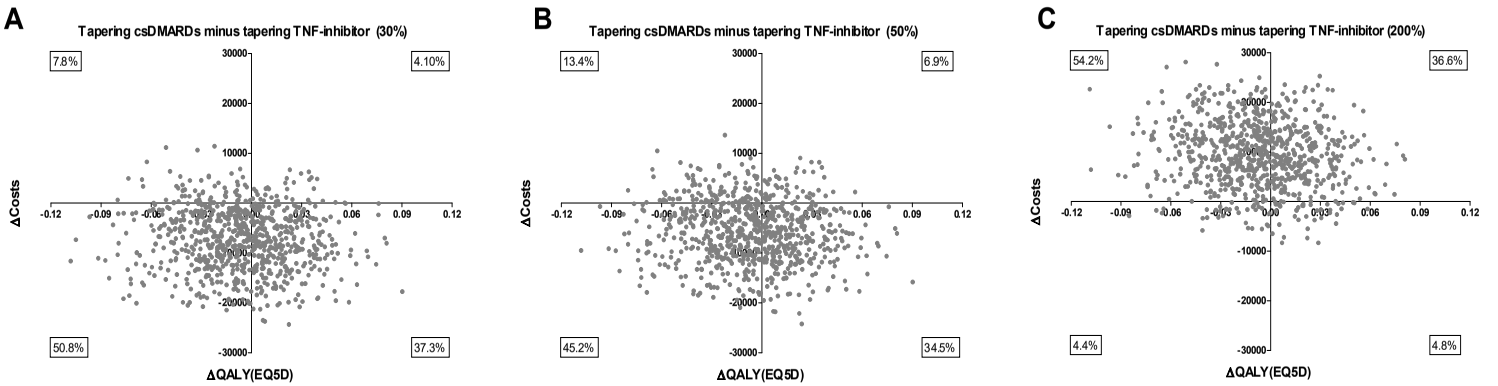


Figure S2 Cost–effectiveness planes with results of sensitivity analyses to investigate biological price changes. (A) 30%, (B) 50%, and (C) 200% of original biological prices were taken to investigate the effect of price changes.

Supplementary table S1 Costs of medication [33]

	Calculation base	Costs
DMARDs		
Methotrexate		
- Oral	2.5mg tablet	€0.18
- Subcutaneous		
• 2.5 tot 10 mg	per piece	€13.72
• 12.5 tot 20 mg	per piece	€27.45
• 22.5 tot 25 mg	per piece	€34.31
Sulfasalazine (oral)	500mg tablet	€0.11
Hydroxychloroquine (oral)	200mg tablet	€0.14
Leflunomide (oral)		
- 20 mg tablet	per piece	€1.48
- 10 mg tablet	per piece	€1.14
Glucocorticoids		
Prednisone (oral)		
- Oral	5mg tablet	€0.05
Triamcinolone (im)	80mg	€6.72
Methylprednisolone (im)	120mg	€6.74
Biologicals		
Etanercept (sc)		
- 25 mg	per piece	€105.63
- 50 mg	per piece	€211.25
Adalimumab (sc)		
- 20 mg	per piece	€269.22
- 40 mg	per piece	€538.44
Certolizumab (sc)		
- 200 mg	per piece	€505.77
Golimumab (sc)		
- 50 mg	per piece	€1125.84
Abatacept (sc)		
- 125 mg	per piece	€277.95

Supplementary table S2 Reference prices for medical consumption [34]

	Reference price
Standard healthcare	
Inpatient day	
- General hospital (day)	€461
- University hospital (day)	€668
Intensive care unit (day)	€1234
Daycare treatment (day)	€495
Outpatient visit	
- Specialist	
• General hospital	€83
• University hospital	€170
Emergency room visit	€269
Primary care physician	€34
Paramedical care	
- Physical therapy	€34
- Occupational therapy	€34
- Speech therapy	€31
- Dietary advice (hour)	€31
Mental healthcare	
- Social worker	€68
- Psychologist	€67
Complementary medicine	
Alternative medical systems	
- Homeopathic treatment	€31

Supplementary table S3 Average hourly productivity costs, stratified for age and sex [35, 36]

Age (years)	Men	Women
15 to 19	€ 8.77	€ 7.97
20 to 24	€ 16.56	€ 16.13
25 to 29	€ 22.53	€ 23.12
30 to 35	€ 28.14	€ 27.81
35 to 40	€ 33.91	€ 30.62
40 to 45	€ 39.13	€ 30.87
45 to 50	€ 42.61	€ 29.67
50 to 55	€ 44.13	€ 28.75
55 to 60	€ 44.48	€ 28.68
60 to 65	€ 43.34	€ 28.83

Supplemental table S4 Comparison of imputed data versus complete cases

	Complete cases	Imputed dataset	Difference	P value
EQ5D, mean (sd)	0.83 (0.16)	0.83 (0.15)	0.002 (0.0053)	0.69
Paid work, proportion (se)	0.506 (0.0019)	0.506 (0.0019)	0.0002 (0.0027)	0.94
LOG days with productivity loss, mean (sd)	1.80 (0.82)	1.80 (0.81)	0.000027 (0.0066)	1
Productivity score, mean (sd)	7.63 (1.81)	7.63 (1.81)	0.000000070	1
LOG number of times absent due to sick leave, mean (sd)	0.26 (0.53)	0.26 (0.53)	0.00000067	1
Duration of sick leave, mean (sd)	2.94 (2.49)	2.94 (2.49)	0.00070	0.97
Working hours per week, mean (sd)	30.3 (11.8)	30.3 (11.7)	0.0015	0.99

EQ5D: EuroQol with 5 dimensions; sd: standard deviation

LITERATURE

1. Klarenbeek NB, Allaart CF, Kerstens PJ, Huizinga TW, Dijkmans BA: The BeSt story: on strategy trials in rheumatoid arthritis. *Curr Opin Rheumatol* 2009, 21(3):291-298.
2. de Jong PH, Hazes JM, Han HK, Huisman M, van Zeben D, van der Lubbe PA, Gerards AH, van Schaeybroeck B, de Sonnaville PB, van Krugten MV *et al*: Randomised comparison of initial triple DMARD therapy with methotrexate monotherapy in combination with low-dose glucocorticoid bridging therapy; 1-year data of the tREACH trial. *Ann Rheum Dis* 2014, 73(7):1331-1339.
3. Smolen JS, Landewe R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, Nam J, Ramiro S, Voshaar M, van Vollenhoven R *et al*: EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017, 76(6):960-977.
4. Biologicals, 2014-2018 [https://www.gipdatabank.nl/]
5. Medicijnkosten [http://www.medicijnkosten.nl]
6. Kuijper TM, Folmer R, Stolk EA, Hazes JMW, Luime JJ: Doctors' preferences in de-escalating DMARDs in rheumatoid arthritis: a discrete choice experiment. *Arthritis Res Ther* 2017, 19(1):78.
7. van Mulligen E, de Jong PHP, Kuijper TM, van der Ven M, Appels C, Bijkerk C, Harbers JB, de Man Y, Molenaar THE, Tchetverikov I *et al*: Gradual tapering TNF inhibitors versus conventional synthetic DMARDs after achieving controlled disease in patients with rheumatoid arthritis: first-year results of the randomised controlled TARA study. *Ann Rheum Dis* 2019, 78(6):746-753.
8. Birkner B, Rech J, Stargardt T: Cost-utility analysis of de-escalating biological disease-modifying anti-rheumatic drugs in patients with rheumatoid arthritis. *PLoS One* 2020, 15(1):e0226754.
9. Kobelt G: Treating to target with etanercept in rheumatoid arthritis: cost-effectiveness of dose reductions when remission is achieved. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2014, 17(5):537-544.
10. Tran-Duy A, Ghiti Moghadam M, Oude Voshaar MAH, Vonkeman HE, Boonen A, Clarke P, McColl G, Ten Klooster PM, Zijlstra TR, Lems WF *et al*: An Economic Evaluation of Stopping versus Continuing TNF-Inhibitor Treatment in Rheumatoid Arthritis Patients in Remission or Low Disease Activity: results from the POET randomized trial. *Arthritis Rheumatol* 2018.
11. Vanier A, Mariette X, Tubach F, Fautrel B, Group SS: Cost-Effectiveness of TNF-Blocker Injection Spacing for Patients with Established Rheumatoid Arthritis in Remission: An Economic Evaluation from the Spacing of TNF-Blocker Injections in Rheumatoid Arthritis Trial. *Value Health* 2017, 20(4):577-585.
12. Kievit W, van Herwaarden N, van den Hoogen FH, van Vollenhoven RF, Bijlsma JW, van den Bemt BJ, van der Maas A, den Broeder AA: Disease activity-guided dose optimisation of adalimumab and etanercept is a cost-effective strategy compared with non-tapering tight control rheumatoid arthritis care: analyses of the DRESS study. *Ann Rheum Dis* 2016, 75(11):1939-1944.
13. Ghiti Moghadam M, Ten Klooster PM, Vonkeman HE, Kneepkens EL, Klaasen R, Stolk JN, Tchetverikov I, Vreugdenhil SA, van Woerkom JM, Goekoop-Ruiterman

YPM *et al*: Impact of Stopping Tumor Necrosis Factor Inhibitors on Rheumatoid Arthritis Patients' Burden of Disease. *Arthritis Care Res (Hoboken)* 2018, 70(4):516-524.

14. van Mulligen E, Weel AE, Hazes JM, van der Helm-van Mil A, de Jong PHP: Tapering towards DMARD-free remission in established rheumatoid arthritis: 2-year results of the TARA trial. *Ann Rheum Dis* 2020.
15. Zwaap J, Knies S, van der Meijden C, Staal P, van der Heiden L: Kosteneffectiviteit in de praktijk. In. Edited by minister van Volksgezondheid WeS. Zorginstituut Nederland: Zorginstituut Nederland; 2015.
16. Cross M, Smith E, Hoy D, Carmona L, Wolfe F, Vos T, Williams B, Gabriel S, Lassere M, Johns N *et al*: The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014, 73(7):1316-1322.
17. Ubel PA, Hirth RA, Chernew ME, Fendrick AM: What is the price of life and why doesn't it increase at the rate of inflation? *Arch Intern Med* 2003, 163(14):1637-1641.
18. Torrance GW: Measurement of health state utilities for economic appraisal. *J Health Econ* 1986, 5(1):1-30.
19. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A: Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *Br J Rheumatol* 1997, 36(5):551-559.
20. Huscher D, Merkesdal S, Thiele K, Zeidler H, Schneider M, Zink A, German Collaborative Arthritis C: Cost of illness in rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis and systemic lupus erythematosus in Germany. *Ann Rheum Dis* 2006, 65(9):1175-1183.
21. Nahin RL, Barnes PM, Stussman BJ, Bloom B: Costs of complementary and alternative medicine (CAM) and frequency of visits to CAM practitioners: United States, 2007. *Natl Health Stat Report* 2009(18):1-14.
22. Hakkaart-van Roijen L, Van der Linden N, Bouwmans CAM, Kanters TA, SS. T: Kostenhandleiding. Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg**.** 2015.
23. Bouwmans C, Krol M, Brouwer W, Severens JL, Koopmanschap MA, Hakkaart L: IMTA Productivity Cost Questionnaire (IPCQ). *Value Health* 2014, 17(7):A550.
24. Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L: The friction cost method for measuring indirect costs of disease. *J Health Econ* 1995, 14(2):171-189.
25. Beloning en arbeidsvolume van werknemers [https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84165NED/table?dl=217EE&ts=1566303305202]
26. Beloningsverschillen tussen mannen en vrouwen [https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81901NED/table?dl=217EC]
27. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL: Methods for the economic evaluation of health care programme. Third edition: Oxford: Oxford University Press; 2005.
28. Graham JW, Olchowski AE, Gilreath TD: How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev Sci* 2007, 8(3):206-213.
29. Pensioenleeftijd werknemers in 2018 [https://www.cbs.nl/nl-nl/maatwerk/2019/32/pensioenleeftijd-werknemers-in-2018]
30. de Jong PH, Hazes JM, Buisman LR, Barendregt PJ, van Zeben D, van der Lubbe PA, Gerards AH, de Jager MH, de Sonnaville PB, Grillet BA *et al*: Best cost-effectiveness and worker productivity with

initial triple DMARD therapy compared with methotrexate monotherapy in early rheumatoid arthritis: cost-utility analysis of the tREACH trial. *Rheumatology (Oxford)* 2016, 55(12):2138-2147.

31. van Herwaarden N, den Broeder AA, Jacobs W, van der Maas A, Bijlsma JW, van Vollenhoven RF, van den Bemt BJ: Down-titration and discontinuation strategies of tumor necrosis factor-blocking agents for rheumatoid arthritis in patients with low disease activity. *Cochrane Database Syst Rev* 2014(9):CD010455.
32. Boonen A, van der Heijde D, Landewe R, Spoorenberg A, Schouten H, Rutten-van Molken M, Guillemin F, Dougados M, Mielants H, de Vlam K *et al*: Work status and productivity costs due to ankylosing spondylitis: comparison of three European countries. *Ann Rheum Dis* 2002, 61(5):429-437.
33. Zorginstituut Nederland. Medicijnkosten. 2019. http://www.medicijnkosten.nl (accessed 18-08-2019 2019).
34. Hakkaart-van Roijen L, Van der Linden N, Bouwmans CAM, Kanters TA, SS. T. Kostenhandleiding. Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg. 2015.
35. Statline C. Beloning en arbeidsvolume van werknemers. 2019. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84165NED/table?dl=217EE&ts=1566303305202 (accessed 01-06-2019 2019).
36. Statline C. Beloningsverschillen tussen mannen en vrouwen. 2019. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81901NED/table?dl=217EC (accessed 01-06-2019 2019).

CHAPTER 10

DISCUSSION

As a result of better treatment outcomes in rheumatoid arthritis (RA), many patients have a well-controlled disease. Continuation of treatment for these patients is not always necessary, and options for tapering of treatment have been previously explored. However, RA remains a chronic, fluctuating, incurable disease. Since it is unclear how and when tapering treatment should take place, it should always be carefully considered.

Within this thesis the following questions concerning tapering of treatment were answered for an established RA population:

- Clinical perspective: How should we taper treatment, and is it possible to discontinue medication completely?
- Patient perspective: What is the impact of a disease flare during tapering on patients' lives?
- Societal perspective: What is the cost-effectiveness of different tapering strategies?

The answers to these questions were addressed using data from the TApering strategies in Rheumatoid Arthritis (TARA) trial.

Clinical perspective

Tapering treatment has been recommended by the EULAR guidelines. They state that when a patient is in persistent remission, after tapering glucocorticoids, rheumatologists could consider tapering bDMARDs, especially when combined with a csDMARD. Also, when a patient is in remission with only a csDMARD, tapering of that csDMARD can be considered.[1] However, no consensus has been reached on the definition of sustained remission, nor is there a pre-specified tapering strategy (i.e. stopping versus gradually tapering).[2] Previous literature has shown that tapering treatment is possible, however, there are no studies directly comparing different tapering strategies.[3-6]

From the results of the TARA trial we concluded that from a clinical perspective the tapering order was not relevant. The flare rates between both tapering strategies, namely tapering the csDMARDs first followed by the TNF-inhibitor or vice versa, were completely similar up to 9 months (chapter 3). A non-significant 10% difference in

flare rates occurred after 12 months, but this difference disappeared in the second year and after two years we found a flare rate of 61% in both tapering arms (chapter 6). Tapering TNF-inhibitors first was, therefore, not superior to tapering csDMARDs first from a clinical perspective.

Nonetheless, a 61% flare rate seems high, but one must keep in mind that these reported flare rates also include flares that would normally occur without tapering. It is known that RA patients who are in sustained remission and are using a stable DMARD dosage can experience a flare. For example, tapering trials with a control arm that continued medication, and thus not tapered, showed flare rates between 12%-48% after one year of follow-up. Flare rates within the tapering arms of aforementioned studies ranged between 51%-77%.[4, 5, 7, 8] In these studies there is an increased risk of flare in the tapering arm, but there is also a risk of flare when medication is not tapered. Flare rates within the TARA trial were comparable with previous tapering trials. These comparisons, however, should be interpreted carefully, because of differences in study design, i.e. differences in patient population, flare definition and starting point of tapering, which resulted in a wide range of flare rates.[9-11]

Besides flare rates, also medication use was taken into account in chapter 3 and 6. Due to the gradual tapering scheme and large amount of tapering steps, six in total, tapering of treatment was stopped at several dosage levels if a flare occurred. After two years, some patients were able to completely stop all DMARDs (15%, chapter 6). On the other hand, only a minority of the patients were not able to taper any medication at all (16%, chapter 6). This indicates that although DMARD-free remission is rare, almost all RA patients can reduce their medication dosages if they have a well-controlled disease. This was also confirmed with real-world data in chapter 7. Future tapering strategies might, therefore, not only be targeted at complete withdrawal, but also at finding the lowest possible DMARD dosage on which patients still have a well-controlled disease. Previous studies investigating reduced dose regimens versus full dose regimens showed non-inferiority, underlining the potential of dose reduction. [4, 7, 11, 12]

If treatment is completely tapered while maintaining remission, patients reach a state of DMARD-free remission, which is the closest to an actual cure for RA nowadays.[13] In chapter 5 we showed that this outcome is achievable for 10%-20% of the whole RA population, and in chapter 6 we showed that this is also achievable in an established

RA population (15% of the TARA patients were in DMARD-free remission). Because only a minority of RA patient are able to reach this goal, DMARD-free remission as a target in a treat-to-target management approach is debatable.

Nowadays, the greatest challenge is to find those patients who are able to (partly) taper medication without flares to prevent the associated negative effects. Current known predictors for flare-free tapering are absence of auto-antibodies and shared epitope (chapter 5). Other (clinical) factors related to flare-free tapering were ambiguous (i.e. symptom duration), or did not show any correlation (i.e. age, erosions at baseline). Within the TARA trial, no relation was found between being in deep remission (DAS<1.6) and the ability to completely taper medication (chapter 6). This implies that current EULAR recommendations concerning eligibility of patients for tapering are too strict, since they recommend to only taper medication when patients are in sustained remission, preferably Boolean based.[1, 2, 14]

Current tapering strategies are still based on a trial-and-error approach, because we cannot adequately predict a disease flare before its occurrence. As a result, high flare rates were seen in all previous tapering trials.[15] Interestingly, in the TARA trial no flares occurred after 18 months, when our patients reached DMARD-free remission, until the end of follow-up, while other studies reported flare rates between 5-25% in the first 6 months after achieving DMARD-free remission.[16-19] This implies that flares in the TARA trial were detected early compared to the previous mentioned trials, which is probably due to the slower and gradual tapering strategy within the TARA trial. Until we find good predictors for flare-free tapering, a stepwise tapering approach with close monitoring still seems to be the safest way to taper medication.

Patient perspective

In the TARA trial we found an overall flare rate of 61% (chapter 6). Of the patients who experienced a flare, 67% regained low disease activity after 6 months. These results were comparable with other tapering studies [3-6, 8]. In all these studies, flare duration was solely based on disease activity measures. The impact of a flare on patients' lives was not taken into account, while it is known that flares can impact daily functioning and can decrease productivity.

Besides the primary outcome (disease flare), secondary outcomes in the TARA trial also included patient reported outcomes (PROs). Comparing the two tapering

strategies, no differences were found in functional ability and quality of life (chapter 3). In chapter 6 adverse events of DMARDs were investigated at baseline, and patients reported a higher burden for use of csDMARDs compared to use of TNF-inhibitors. This might demonstrate that patients have a preference for tapering csDMARDs over TNF-inhibitors.

The impact of a flare on TARA patients was objectified by quantifying changes in PROs (chapter 8). We showed that a disease flare has a significant effect on all components of the disease activity score, and also on functional ability, quality of life and fatigue. The largest PRO effects were found in the physical outcomes, which were almost all above the minimal clinically important difference (MCID) and lasted for more than six months. Most importantly, the effect of a flare on patient relevant outcomes could still be observed after the disease activity normalized. This implies that the effect of flare can be detected for a longer period of time than what we now measure with disease activity only. Furthermore, all analyses were performed on a group level, so the effect on the individual patient might be even worse. The true effect of a flare on a patient, therefore, cannot only be visualized with a disease activity measure. Possibly PROs could be monitored more often in daily practice.

Moreover, some of the investigated PROs already worsened before the actual flare occurred, i.e. functional ability and severity of morning stiffness, which might be useful for flare prediction during tapering. These changes before the actual flare occurred were all non-significant, which was also found by a study on prediction on flares by changes in PROs.[20] However, within this study a disease flare was always accompanied with worsening of the PROs, and results from the TARA still indicate that patients already have more complaints before the actual flare was objectified by the treating physician. Due to the long-lasting effect of a flare on patients' lives, it is of great importance that we find good predictors for flare-free tapering. By preventing flares before its occurrence, we automatically prevent the impact of the flare itself. Since we are unable to predict a flare, close monitoring by using PROs seems to be valuable to assess patients' wellbeing before, during, and also after a flare.

Societal perspective

An important reason for tapering medication is reduction of health care costs. On the short term, tapering medication can reduce medication costs, but also long term costs should be taken into account, such as societal costs. Societal costs are costs

due to loss of productivity. In a cost-utility analysis it is important to take into account all associated costs, thus health care costs as well as societal costs.[21]

Cost-effectiveness of tapering csDMARDs or TNF-inhibitors first was investigated in chapter 9. Previous studies investigated the cost-effectiveness of tapering versus continuation of medication. Tapering treatment resulted in cost-savings and a small loss of quality adjusted life years (QALYs).[22, 23] However, these studies were not consistent in their societal costs calculation; often productivity loss was omitted, therefore, the true effect on society remained unclear.

Within the TARA trial we found significant higher medication costs when the csDMARDs were tapered first, however this was counterbalanced by higher societal costs due to more productivity loss in the group who tapered the TNF-inhibitor first (chapter 9). It is known that decreasing costs for bDMARDs might influence outcomes of cost-effectiveness analyses. In chapter 9 the impact of changes in prices of medication was, therefore, investigated. In case of higher bDMARDs costs, the strategy in which the TNF-inhibitors were tapered first appeared to be more cost-effective. If, on the other hand, prices of bDMARDs decreased, the strategy in which csDMARDs were tapered first was more cost-effective. Future prices of bDMARDs will probably decline further, because of the introduction of biosimilars.[24] Therefore, I expect that tapering of csDMARDs first will become more cost-effective compared to tapering the TNF-inhibitor first.

In the TARA trial, the societal costs were more than half of the total costs and had a major impact on the outcome. So, despite the differences in medication costs, no difference in overall cost-effectiveness were found between both tapering strategies. This underlines the importance of including all costs, including those caused due to productivity loss, to prevent that tapering decisions are solely based on medication costs.

Previous studies showed that savings on health care and societal costs could be obtained by treating to target within newly diagnosed early RA patients.[25] Ideally, if RA patients are in sustained remission, medication is quickly tapered and possibly stopped to reduce health care costs. On the short term, this requires regular monitoring of patients who want to taper their medication, which could increase the costs for medical consumption. Possibly, future monitoring might be also possible with PROs, which could also decrease these costs because this can be done at home.

On the long term, some patients will be able to reach DMARD-free remission. This means that these patients neither require medication nor do they need regular visits to the outpatient clinic, which leads to reductions in health care costs.

General considerations

The TARA study is a single-blinded randomized controlled trial, which has been conducted in 13 hospitals in the South-Western part of the Netherlands. The study was initiated in 2011 and finished in 2018. The TARA trial was designed without a control arm, which might seem odd for a randomized controlled trial. However, previous research already showed that tapering treatment is possible in RA patients with a well-controlled disease. The TARA trial was, therefore, the logical next step by comparing different tapering strategies.

A general limitation of the TARA trial was that, due to difficulties with inclusions, the targeted sample size was not reached. Inclusion criteria at the start of the TARA trial were a DAS ≤ 1.6 and a maximum of one swollen joint, and use of methotrexate and either etanercept or adalimumab. These inclusion criteria were broadened to a DAS ≤ 2.4 and a maximum of one swollen joint, use of csDMARDs combined with etanercept, adalimumab, certolizumab or golimumab. Despite these changes, the inclusion difficulties remained, and eventually led to preliminary termination of inclusion in December 2017. Instead of the targeted 208 patients, which included a drop-out percentage of 10%, 189 patients were included. After one year of follow-up, we encountered a slightly lower drop-out rate of 8%. For the primary outcome we performed a sensitivity analysis in chapter 3, and our results appeared to be valid.

Except for chapters 2 and 5, all chapters were based on data derived from the TARA trial. The primary outcome (disease flares after one year of follow-up) was investigated in chapter 3. Other outcomes were secondary (chapter 6, 7, 9), including one post-hoc analysis (chapter 8). In chapter 2 real-world data was used from a retrospective cohort study.

The TARA trial has several strengths, including being a randomized controlled trial. Advantages of randomized controlled trials are minimizing confounding and selection bias due to randomization, and information bias due to blinding.

Selection bias occurs when the patient sample is not representative for the targeted

population, and can affect generalizability of outcomes. For the TARA trial patients were selected who were willing to try to taper their medication, and our result will therefore be generalizable to established RA patients who are willing to taper their medication. Some selection bias might have occurred due to the preference of rheumatologists to recruit patients based on their opinion on feasibility of tapering. However, based on the baseline characteristics (chapter 3, 4) the TARA population appeared to be a reflection of a normal, established RA population.

Information bias leads to errors in results due to misclassification or measurement errors. Information bias can be non-differential or differential. Non-differential information bias occurs due to random measurement errors in outcomes and will lead to dilution of the effect towards the null, which we cannot correct for. Differential information bias is related to the outcome, and can lead, in case of the TARA trial, to overestimation of the effect. Differential information bias can be minimized by blinding of outcome assessment. Within the single-blinded TARA trial, the research nurses, who performed the DAS assessment, were blinded to the allocated treatment arm. Treating rheumatologists and patients were aware of the allocated treatment arm. Their believes on tapering of treatment might have influenced decisions to intensify or taper treatment. Based on the reported adverse events (chapter 7) I suspect that patients preferred to taper the csDMARD, also because all participating patients previously had to intensify their treatment with a bDMARD to reach well-controlled disease. Most likely, both rheumatologists and patients therefore expected that tapering of the bDMARDs would lead to more flares than tapering csDMARDs, in which the effect could be overestimated. Since we did not observe a significant difference in flare rates between tapering arms, I believe our outcomes are valid.

Confounding will occur when there is an external factor (confounder) that has an association between the factor of interest and the outcome. If one does not correct for confounding, the results may be distorted. Due to randomization, confounding was minimized. Within the TARA trial minimization randomization was used, which aims at reducing imbalances between patients in each group, taking into account certain predefined factors. Normally, randomization leads to a balanced allocation of patients over two arms. Because of 13 participating centers at different locations with different non-blinded rheumatologists (with different believes), it was a concern that when allocation would not be balanced within the participating centers, comparison between tapering arms would suffer from bias due to preferences of rheumatologists

and/or hospital specific treatment strategies for RA patients. Therefore, minimization randomization was stratified and balanced for each participating center. As a result, we also had to correct for this in all analyses. On the other hand, by stratifying the randomization, potential confounding could be introduced. To investigate this, the corrected outcomes were compared to the uncorrected outcomes, and no differences were observed. This indicates that no confounding was introduced by stratifying the randomization.

In chapter 8 a post hoc analysis was performed and new groups were made (i.e. flare versus non-flare patients). A major drawback of post hoc analyses are the risk on 'data dredging'. However, in this case the research question was created before the data was analyzed and corrections for multiple testing were made. As far as I know, no confounding factors were present.

All perspectives combined: personal recommendations for clinical practice

Disease activity guided tapering of DMARDs has proven to be feasible, safe and effective in RA patients with low disease activity or remission, but is known to be accompanied with a higher risk of flares. Taking into account the clinical perspective, the patient perspective, and the societal perspective, I have the following recommendations for clinical practice:

- In the TARA trial it was shown that almost all established RA patients with a sustained, well-controlled disease, defined as a DAS<2.4 and no swollen joints, were able to taper medication. Therefore I suggest that tapering could be initiated when patients have a DAS ≤ 2.4 and no swollen joints for at least 6 months.

- The order of tapering, i.e. csDMARDs or bDMARDs first, is not relevant based on both clinical outcomes (flare rates), as well as the cost-effectiveness analysis. Tapering treatment should always be carefully considered, and should always be based on a shared decision between the rheumatologist and the patient. Given the facts that (1) patients experience more burden from csDMARDs usage compared to TNF-inhibitor usage, (2) the majority of patients is able to taper part of their medication, especially in the group who tapered csDMARDs first, and (3) prices of bDMARDs will probably decline in the future, I would propose to taper csDMARDs first.

- Tapering treatment is associated with an increased risk of flares. Therefore, my advice is to gradually taper medication and closely monitor patients during tapering to detect a possible disease flare early, and intensify medication if necessary. Monitoring could be improved by not only focusing on measuring disease activity, but also on patients' wellbeing. Measuring PROs on a regular basis could help rheumatologists in identifying a flare early. Since the effect of a flare is still apparent after disease activity normalizes, monitoring PROs might help evaluate patients' wellbeing.

Future research suggestions

- Tapering strategies

- Although the majority of the TARA patients were not able to completely taper all medication, and thus did not reach DMARD-free remission, almost all TARA patients were capable of tapering some of their medication. Therefore, future research should focus on finding a more tailor-made tapering approach, in which the ultimate aim is to reach the lowest medication dosage at which patients still have a well-controlled disease.

- The flare rates in the TARA trial were comparable with other studies in which no gradual approach was used. In total, it took 18 months to completely taper medication. Because of the similar flare rates, those 18 months might not be necessary for gradual tapering. Future research could explore whether if it is required to taper DMARDs separately and one by one, while possibly multiple DMARDs can be tapered at once or tapering can be alternated between DMARDs. Also tapering of other DMARDs (i.e. bDMARDs with another mode of action) could be explored. Furthermore, shared decision making can be used to determine how tapering should take place, as well as which DMARD(s) will be tapered.

- Within this thesis we highlighted several aspects of tapering from different viewpoints. Still, the decision on when to commence tapering of treatment remains unanswered. Often it is suggested that tapering can only be initiated when patients reach 'deep' or 'true' remission, however a debate is ongoing about what the true definition of remission

is. If a stricter definition of remission is used, less patients are eligible to taper medication. However, we have already shown that tapering is also possible for patients with a well-controlled disease, defined as a DAS ≤ 2.4 without any swollen joints. Future research should, therefore, focus more on criteria that make flare-free tapering possible, thereby not only focusing on the currently existing remission criteria.

- **Prediction of flares**

  - The best study design for comparing treatment strategies is a randomized controlled trial. However, trials are known to have limited generalizability and usually have a relative short follow-up period. For the prediction of flares, I believe a real-world cohort is the best way to assess factors influencing flare-free survival. Besides results from clinical trials, it is important to obtain more real-world information to improve generalizability.

  - Future research should focus on finding biomarkers that predict who is eligible for tapering treatment, and, following this, to identify factors that predict flare-free survival. If patients can be stratified according to their chances at flare-free tapering of treatment, tapering can be initiated. Possible predictors should also include patient reported outcomes, because we have already shown that they worsen before an actual flare occurs.

- **DMARD-free remission**

  - Future research on tapering should focus more on the long term, in which also DMARD-free remission should be assessed at least six months after complete withdrawal of medication. A follow-up of only one year is too short to correctly evaluate DMARD-free remission rates. Since recruitment of patients for tapering trials has shown to be difficult, and a trial with a long follow-up is costly, this research could be performed in large registries in which tapering occurs. I expect that tapering of treatment will occur more often in the future due to the incorporation of DMARD tapering in the guidelines and the growing knowledge on tapering of treatment in general.

Literature

1. Smolen JS, Landewe RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, McInnes IB, Sepriano A, van Vollenhoven RF, de Wit M *et al*: EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. *Ann Rheum Dis* 2020.
2. Studenic P, Felson D, de Wit M, Alasti F, Stamm TA, Smolen JS, Aletaha D: Testing different thresholds for patient global assessment in defining remission for rheumatoid arthritis: are the current ACR/EULAR Boolean criteria optimal**?** *Ann Rheum Dis* 2020, 79(4):445-452.
3. Smolen JS, Nash P, Durez P, Hall S, Ilivanova E, Irazoque-Palazuelos F, Miranda P, Park MC, Pavelka K, Pedersen R *et al*: Maintenance, reduction, or withdrawal of etanercept after treatment with etanercept and methotrexate in patients with moderate rheumatoid arthritis (PRESERVE): a randomised controlled trial. *Lancet* 2013, 381(9870):918-929.
4. van Herwaarden N, van der Maas A, Minten MJ, van den Hoogen FH, Kievit W, van Vollenhoven RF, Bijlsma JW, van den Bemt BJ, den Broeder AA: Disease activity guided dose reduction and withdrawal of adalimumab or etanercept compared with usual care in rheumatoid arthritis: open label, randomised controlled, non-inferiority trial. *Bmj* 2015, 350:h1389.
5. Fautrel B, Pham T, Alfaiate T, Gandjbakhch F, Foltz V, Morel J, Dernis E, Gaudin P, Brocq O, Solau-Gervais E *et al*: Step-down strategy of spacing TNF-blocker injections for established rheumatoid arthritis in remission: results of the multicentre non-inferiority randomised open-label controlled trial (STRASS: Spacing of TNF-blocker injections in Rheumatoid ArthritiS Study). *Ann Rheum Dis* 2016, 75(1):59-67.
6. Haschka J, Englbrecht M, Hueber AJ, Manger B, Kleyer A, Reiser M, Finzel S, Tony H-P, Kleinert S, Feuchtenberger 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* 2016, 75(1):45-51.
7. van Vollenhoven RF, Ostergaard M, Leirisalo-Repo M, Uhlig T, Jansson M, Larsson E, Brock F, Franck-Larsson K: Full dose, reduced dose or discontinuation of etanercept in rheumatoid arthritis. *Ann Rheum Dis* 2016, 75(1):52-58.
8. Ghiti Moghadam M, Vonkeman HE, Ten Klooster PM, Tekstra J, van Schaardenburg D, Starmans-Kool M, Brouwer E, Bos R, Lems WF, Colin EM *et al*: Stopping Tumor Necrosis Factor Inhibitor Treatment in Patients With Established Rheumatoid Arthritis in Remission or With Stable Low Disease Activity: A Pragmatic Multicenter, Open-Label Randomized Controlled Trial. *Arthritis Rheumatol* 2016, 68(8):1810-1817.
9. Kuijper TM, Lamers-Karnebeek FB, Jacobs JW, Hazes JM, Luime JJ: Flare Rate in Patients with Rheumatoid Arthritis in Low Disease Activity or Remission When Tapering or Stopping Synthetic or Biologic DMARD: A Systematic Review. *J Rheumatol* 2015, 42(11):2012-2022.
10. Edwards CJ, Fautrel B, Schulze-Koops H, Huizinga TWJ, Kruger K: Dosing down with biologic therapies: a systematic review and clinicians' perspective. *Rheumatology (Oxford)* 2017, 56(11):1847-1856.
11. Verhoef LM, van den Bemt BJ, van der Maas A, Vriezekolk JE, Hulscher ME, van den Hoogen FH, Jacobs WC, van Her-

waarden N, den Broeder AA: Down-titration and discontinuation strategies of tumour necrosis factor-blocking agents for rheumatoid arthritis in patients with low disease activity. *Cochrane Database Syst Rev* 2019, 5:CD010455.

12. Kavanaugh A, Lee SJ, Curtis JR, Greenberg JD, Kremer JM, Soto L, Etzel CJ, Cox V, Yoshida K, Reed GW *et al*: Discontinuation of tumour necrosis factor inhibitors in patients with rheumatoid arthritis in low-disease activity: persistent benefits. Data from the Corrona registry. *Ann Rheum Dis* 2015, 74(6):1150-1155.
13. Boeters DM, Burgers LE, Toes RE, van der Helm-van Mil A: Does immunological remission, defined as disappearance of autoantibodies, occur with current treatment strategies? A long-term follow-up study in rheumatoid arthritis patients who achieved sustained DMARD-free status. *Ann Rheum Dis* 2019, 78(11):1497-1504.
14. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, Aletaha D, Allaart CF, Bathon J, Bombardieri S et al: American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann Rheum Dis* 2011, 70(3):404-413.
15. Schett G, Emery P, Tanaka Y, Burmester G, Pisetsky DS, Naredo E, Fautrel B, van Vollenhoven R: Tapering biologic and conventional DMARD therapy in rheumatoid arthritis: current evidence and future directions. *Ann Rheum Dis* 2016, 75(8):1428-1437.
16. Kuijper TM, Luime JJ, de Jong PH, Gerards AH, van Zeben D, Tchetverikov I, de Sonnaville PB, van Krugten MV, Grillet BA, Hazes JM *et al*: Tapering conventional synthetic DMARDs in patients with early arthritis in sustained remission: 2-year follow-up of the tREACH trial. *Ann Rheum Dis* 2016, 75(12):2119-2123.
17. Emery P, Burmester GR, Bykerk VP, Combe BG, Furst DE, Barre E, Karyekar CS, Wong DA, Huizinga TW: Evaluating drug-free remission with abatacept in early rheumatoid arthritis: results from the phase 3b, multicentre, randomised, active-controlled AVERT study of 24 months, with a 12-month, double-blind treatment period. *Ann Rheum Dis* 2015, 74(1):19-26.
18. Bijlsma JWJ, Welsing PMJ, Woodworth TG, Middelink LM, Petho-Schramm A, Bernasconi C, Borm MEA, Wortel CH, Ter Borg EJ, Jahangier ZN *et al*: Early rheumatoid arthritis treated with tocilizumab, methotrexate, or their combination (U-Act-Early): a multicentre, randomised, double-blind, double-dummy, strategy trial. *Lancet* 2016, 388(10042):343-355.
19. Klarenbeek NB, van der Kooij SM, Guler-Yuksel M, van Groenendael JH, Han KH, Kerstens PJ, Huizinga TW, Dijkmans BA, Allaart CF: Discontinuing treatment in patients with rheumatoid arthritis in sustained clinical remission: exploratory analyses from the BeSt study. *Ann Rheum Dis* 2011, 70(2):315-319.
20. Walter MJ, Mohd Din SH, Hazes JM, Lesaffre E, Barendregt PJ, Luime JJ: Is tightly controlled disease activity possible with online patient-reported outcomes? *J Rheumatol* 2014, 41(4):640-647.
21. Alemao E, Al MJ, Boonen AA, Stevenson MD, Verstappen SMM, Michaud K, Weinblatt ME, Rutten-van Molken M: Conceptual model for the health technology assessment of current and novel interventions in rheumatoid arthritis. *PLoS One* 2018, 13(10):e0205013.
22. Vanier A, Mariette X, Tubach F, Fautrel B, Group SS: Cost-Effectiveness of TNF-Blocker Injection Spacing for Patients with Established Rheumatoid Arthritis in Remission: An Economic Evaluation from the Spacing of TNF-Blocker Injections in Rheumatoid Arthritis Trial. *Value Health* 2017, 20(4):577-585.

23. Kievit W, van Herwaarden N, van den Hoogen FH, van Vollenhoven RF, Bijlsma JW, van den Bemt BJ, van der Maas A, den Broeder AA: Disease activity-guided dose optimisation of adalimumab and etanercept is a cost-effective strategy compared with non-tapering tight control rheumatoid arthritis care: analyses of the DRESS study. *Ann Rheum Dis* 2016, 75(11):1939-1944.
24. Kim H, Alten R, Avedano L, Dignass A, Gomollón F, Greveson K, Halfvarson J, Irving PM, Jahnsen J, Lakatos PL *et al*: The Future of Biosimilars: Maximizing Benefits Across Immune-Mediated Inflammatory Diseases. *Drugs* 2020, 80(2):99-113.
25. de Jong PH, Hazes JM, Buisman LR, Barendregt PJ, van Zeben D, van der Lubbe PA, Gerards AH, de Jager MH, de Sonnaville PB, Grillet BA *et al*: Best cost-effectiveness and worker productivity with initial triple DMARD therapy compared with methotrexate monotherapy in early rheumatoid arthritis: cost-utility analysis of the tREACH trial. *Rheumatology (Oxford)* 2016, 55(12):2138-2147.

ADDENDUM

SUMMARY

Treatment outcomes in rheumatoid arthritis (RA) have improved enormously in the last decades due to early initiation of therapy, a treat-to-target approach, and use of biological disease-modifying anti-rheumatic drugs (bDMARDs). As a result, remission in RA occurs more often. This has raised the question whether we need to continue or taper treatment. Reasons for tapering are reduction in costs, since treatment with bDMARDs is very expensive, prevention of possible long term side effects and patient preference. However, by tapering medication the risk of having a disease flare increases, which can have a great impact on patients' lives and on society due to productivity loss, i.e. sick leave or unemployment. Therefore, before tapering treatment is considered, it is important to consider these different viewpoints before making a (shared) decision. Moreover, the optimal tapering strategy leading to the least amount of flares has not been developed yet.

Therefore, the aim of this thesis was to study the clinical-, patient-, and the societal perspective of tapering to help rheumatologist and RA patients with their decisions whether they should taper or not. These perspectives were addressed using data from the TApering strategies in Rheumatoid Arthritis (TARA) trial. The TARA trial was set-up to investigate the best tapering strategy in RA patients with a well-controlled disease. Adult RA patients with a disease activity score (DAS) ≤2.4 and a swollen joint count (SJC) ≤1 for more than three months, using a combination of a csDMARD and TNF-inhibitor, were included. Patients were randomised into gradual tapering either the csDMARD in the first year followed by the TNF-inhibitor in the second year, or vice versa. csDMARD tapering was realized by cutting the dosage into half, a quarter and thereafter it was stopped. The TNF-inhibitor was tapered by doubling the dose interval, followed by cutting the dosage into half, and thereafter it was stopped. The total tapering schedule took 6 months, with dose adjustments every 3 months as long as there was still a well-controlled disease. If a disease flare occurred, defined as DAS>2.4 and/or SJC>1, tapering was stopped and the last effective treatment was restarted and if necessary, medication was intensified further according to a treat-to-target approach, until low disease activity was reached again. After a flare, no further attempts were taken to taper medication during the remainder of the study.

To objectify the possible impact of tapering treatment in daily clinical practice, we described current biological use in rheumatoid arthritis patients in the Netherlands

in **chapter 2**. We stratified real-world biological survival data for discontinuation reason and determined its influenceability. Discontinuation reasons for the first-line biological were mainly ineffectiveness, adverse events, or remission. Biological survival diminished with the number of biologicals used. Biological survival was prolonged if patients had concomitant use of csDMARDs. Rheumatoid factor and Anti-Citrullinated Protein Antibody were negatively associated with respectively biological survival and discontinuation due to remission. Therefore, tailoring treatment based upon autoantibody status might be the first step towards personalized medicine in RA.

In **chapter 3** the first year results of the TARA trial were presented. The primary outcome was the difference in flare rates during the first year of follow-up, between two gradual tapering strategies, namely tapering the TNF-inhibitor of the csDMARD. Up to 9 months, flare rates of tapering csDMARDs or TNF-inhibitors were completely similar. After one year, a non-significant 10% difference was found, favoring csDMARD tapering. Also, no differences were found in disease activity, functional ability, and quality of life. Therefore, tapering the TNF-inhibitors first was not superior to tapering the csDMARDs first and thus from a clinical viewpoint it does not matter which medication is tapered first.

In **chapter 4** more information was given on the treatment strategies that included RA patients in the TARA trial were using. Half of the patients used more than one csDMARD at baseline. Furthermore, an additional sub analysis was performed in which patients using oral glucocorticoids were excluded, whereafter the flare rates between both tapering strategies were compared again. This sub analysis showed similar results and, therefore, the conclusion drawn in chapter 3 is still valid.

If treatment is completely tapered, a patient will be in DMARD-free remission (DFR). In **chapter 5** we performed a systematic literature review investigating the feasibility of DFR as a novel and sustainable outcome for RA. DFR appeared to be achievable in RA, and is sustainable in 10%-20% of patients. However, flares occurred frequently during DMARD-tapering and in the first year after achieving DFR. Although absence of auto-antibodies and shared epitope alleles increased the chance of achieving DMARD-free remission, many other (known) risk factors/patient characteristics lacked association with DFR. DFR can become an important outcome measure for clinical trials, and requires consistency in the definition. Considering the high rate of flares in the first year after DMARD-stop, a DMARD-free follow-up of >12 months is advisable

to evaluate sustainability.

Following this, in **chapter 6** we investigated whether DMARD-free remission is also an achievable treatment outcome in an established RA population using data from the TARA trial. Also, the two-year clinical effectiveness of the two gradual tapering strategies in the TARA trial was assessed. DMARD-free remission was achievable in 15% of established RA patients, and was slightly more frequent in patients who tapered their csDMARD first. Although DMARD-free remission occurs less frequent, most of the RA patients with a well-controlled disease could lower their DMARD dosages. To illustrate, 83% of the patients were able to reduce their medication dosages. However, the tapering order did not influence aforementioned results. Moreover, no difference in flare rates, disease activity, functional ability, or radiographic progression were seen after two years and over time. Because of similar effects from a clinical viewpoint, financial arguments may influence the decision to taper TNF-inhibitors first. In **chapter 7** we compared the TARA data with data from a real-world cohort from Pakistan, which confirmed that the majority of RA patients are able to gradually taper DMARDs.

In **chapter 8** the impact of a disease flare on patient's lives was determined. This was investigated by measuring patient reported outcomes (PROs) during tapering of treatment and comparing the PRO norm values, defined as the average of each PRO 12, 9 and 6 months prior to a flare, with the values at the moment of flare, 3 months prior to a flare and every 4 months after a flare. A flare negatively influenced general health, morning stiffness, functional ability, quality of life, and fatigue and this effect lasted for 6 months. The effect sizes exceeded the minimum clinically important difference for the specific outcome measure, except for functional ability. Although on a group level effect sizes for the separate PROs were not always significant or larger than specific MCIDs, a disease flare can still be of great importance for individual patients.

An important reason for tapering treatment is to save costs. In **chapter 9** we, therefore, investigated the societal impact of tapering. For this, the two year cost-utility ratio between both tapering strategies in the TARA trial was evaluated. Health care costs were lower in the patients who tapered the TNF-inhibitor first, while costs due to productivity loss were higher. Overall, total costs did not differ between both tapering strategies. Regardless of the willingness-to-pay (WTP) threshold, tapering either the TNF-inhibitor or the csDMARD first was equally cost-effective.

In **chapter 10** a general discussion of the main findings of this thesis are provided. New insights from findings within this thesis and their clinical implications are discussed. Also methodological considerations are discussed and their possible implications for the results. Finally, recommendations for clinical practice and suggestions for future research are presented. The recommendations are that tapering can be commenced as soon as patients have a DAS<2.4 for at least six months, without swollen joints. The order of tapering is not relevant, however because patients reported more burden for use of csDMARDs, most patients are able to taper some medication, especially in the group who tapered csDMARDs first, and prices of bDMARDs will decline in the future, I recommend to taper the csDMARD first. Furthermore, to reduce the impact of flares, patients should follow a gradual tapering protocol with close monitoring, preferably including PROs to evaluate their wellbeing.

SAMENVATTING

Behandeluitkomsten bij reumatoïde artritis (RA) zijn de afgelopen jaren sterk verbeterd door vroegdiagnostiek met snelle initiatie van therapie, doelgerichte behandeling en het gebruik van biologicals. Patiënten hebben hierdoor steeds vaker een goed gecontroleerde ziekte of zelfs remissie, waarbij geen gezwollen of pijnlijke gewrichten waarneembaar zijn. Logischerwijs kan men zich dus afvragen of de behandeling voortgezet dient te worden bij deze groep RA patiënten die (langdurig) in remissie zijn. Andere voordelen van afbouwen van medicatie zijn besparing van kosten, het voorkomen van bijwerkingen bij (langdurig) medicatiegebruik en de wens van de patiënt. Een groot nadeel van afbouwen is dat de reuma weer kan opvlammen, wat een groot effect kan hebben op het leven van de patiënt, maar ook op de samenleving door verlies van productiviteit door ziekteverzuim of werkloosheid. De keuze om medicatie af te bouwen dient daarom vanuit verschillende perspectieven worden belicht. Daarnaast is er nog veel onduidelijkheid over wat de optimale afbouwstrategie is.

In dit proefschrift hebben worden twee verschillende afbouwstrategieën vanuit een klinisch-, patiënten-, en maatschappelijk perspectief met elkaar vergeleken. Het uiteindelijke doel was om de reumatoloog beter te informeren over de positieve en negatieve gevolgen van het afbouwen van medicatie om zodoende beter tot een gezamenlijke beslissing te komen. Alle informatie komt uit de Tapering strategies in Rheumatoid Arthritis (TARA) trial. In dit gerandomiseerd onderzoek werden twee afbouwstrategieën met elkaar vergeleken, namelijk het eerst afbouwen van de conventionele reuma-remmers (csDMARDs) gevolgd door de TNF-inhibitor, of eerst afbouwen van de TNF-inhibitor gevolgd door de csDMARD. Deelnemers aan dit onderzoek waren RA patiënten met een rustige ziekte, gedefinieerd als een DAS ≤2.4 en maximaal één gezwollen gewricht gedurende ten minste drie maanden. Voor het afbouwen van de csDMARD werd eerst de dosis gehalveerd, na 3 maanden werd dit nogmaals gehalveerd en na 6 maanden werd de csDMARD gestopt. De TNF-inhibitor werd afgebouwd door eerst het interval tussen twee giften te verdubbelen, vervolgens werd de dosis gehalveerd en daarna werd het medicijn gestopt. Het afbouwen van één van beide medicijnen nam zes maanden in beslag, met elke drie maanden een dosisaanpassing zo lang de ziekte rustig was. Per jaar werd één medicijn afgebouwd. Indien er een ziekteopvlamming was, gedefinieerd als DAS>2.4 of meer dan één gezwollen gewricht, werd de medicatie opgehoogd totdat het behandeldoel, lage

ziekteactiviteit, weer werd bereikt. Na een opvlamming werd tijdens de gehele studie niet weer geprobeerd af te bouwen.

Om de impact van het afbouwen van medicatie bij RA patiënten in de dagelijkse praktijk in kaart te brengen, beschrijven we het gebruik van biologicals in Nederland in **hoofdstuk 2**. We stratificeerden de stopredenen van biologicals en bepaalden hoe dit eventueel beïnvloed kan worden. Redenen om biologicals te staken waren voornamelijk ineffectiviteit, bijwerkingen of remissie. Des te meer biologicals een patiënt gebruikt had, des te korter de gebruiksduur. De gebruiksduur was echter langer wanneer een biological met een csDMARD werd gecombineerd. Verder was aanwezigheid van reumafactor negatief geassocieerd met gebruiksduur en aanwezigheid van en anti-CCP negatief geassocieerd met de mogelijkheid om medicatie af te bouwen. In de toekomst kan behandeling mogelijk meer op maat worden gegeven indien rekening wordt gehouden met de aan- of afwezigheid van auto-antistoffen.

In **hoofdstuk 3** worden de resultaten van het eerste jaar van de TARA trial gepresenteerd vanuit een klinisch perspectief. De primaire uitkomstmaat was het procentuele verschil in ziekte-opvlammingen tussen de twee afbouwstrategieën. Tot negen maanden was er geen enkel verschil in het aantal ziekteopvlammingen tussen beide afbouwstrategieën. Echter, na 12 maanden hadden patiënten die eerst hun TNF-inhibitor afbouwden 10% meer kans (niet-significant) op een ziekte-opvlamming ten opzichte van de patiënten die eerst hun csDMARD afbouwden. Verder werden geen verschillen gevonden in ziekteactiviteit, functioneren of kwaliteit van leven na één jaar en over de tijd. Vanuit een klinisch oogpunt maakt het dus niet uit welk medicament als eerst wordt afgebouwd.

In **hoofdstuk 4** wordt meer informatie gegeven over welke csDMARDs werden gebruikt door patiënten bij inclusie in de TARA studie. Het merendeel van de patiënten gebruikte meer dan één csDMARD naast hun TNF-inhibitor. Daarnaast werd een sub-analyse verricht met weer als uitkomst het procentuele verschil in ziekte-opvlammingen tussen beide afbouwstrategieën, maar waarbij patiënten die glucocorticosteroïden gebruikten werden uitgesloten. Dit leverde dezelfde resultaten op als hoofdstuk 3, waardoor onze conclusie valide bleek.

Indien alle medicatie volledig is afgebouwd, komt de patiënt in DMARD-vrije remissie, wat op dit moment het dichtst in de buurt van genezing van RA ligt. In **hoofdstuk 5**

wordt een systematisch literatuuronderzoek uitgevoerd waarin gekeken wordt naar de haalbaarheid van DMARD-vrije remissie. Ongeveer 10-20% van de RA patiënten kan DMARD-vrije remissie bereiken. Voornamelijk bij het afbouwen van medicatie en in het eerste jaar na het stoppen van de medicatie kregen veel patiënten een ziekteopvlamming. Alleen de afwezigheid van auto-antistoffen en shared epitope allelles verhoogde de kans op het bereiken van DMARD-vrije remissie, terwijl andere (bekende) risicofactoren niet geassocieerd waren met het bereiken van DMARD-vrije remissie. DMARD-vrije remissie kan in de toekomst een belangrijke meetuitkomst worden voor onderzoek, maar hiervoor is wel een eenduidige definitie van DMARD-vrije remissie nodig. Door de hoge aantallen opvlammingen in het eerste jaar na stoppen van DMARDs, is een periode van minimaal 12 maanden na het stoppen van de DMARDs nodig om het als echte DMARD-vrije remissie te kunnen definiëren.

Vervolgens onderzochten we in **hoofdstuk 6** of DMARD-vrije remissie ook een haalbare behandeluitkomst is voor patiënten met een langer bestaande RA. Wederom werd hiervoor data van de TARA studie gebruikt, maar nu werd gekeken over een periode van twee jaar. Slechts 15% van de geïncludeerde RA patiënten was in DMARD-vrije remissie en dit was niet afhankelijk van de afbouwstrategie. Hoewel een klein percentage van de RA patiënten na twee jaar in DMARD-vrije remissie was, kon het overgrote gedeelte (83%) van de geïncludeerde RA patiënten een deel van hun medicatie afbouwen. Verder constateerde we wederom dat de volgorde van medicatie afbouwen geen invloed had op het aantal opvlammingen, ziekteactiviteit, functioneren en radiologische progressie. In **hoofdstuk 7** worden onze klinische uitkomsten vergeleken met de dagelijkse reumatologische zorg in Pakistan. Onze resultaten waren vergelijkbaar met de aangedragen real-world data, het blijkt dat ook in de dagelijkse praktijk de meerderheid van de RA patiënten in staat is om medicatie gradueel af te bouwen.

In **hoofdstuk 8** werd onderzocht wat de impact van een ziekteopvlamming is op het leven van de patiënt. In de TARA studie werden frequent patiënt relevante uitkomsten (PROs) gemeten. Deze PROs werden genormeerd door het gemiddelde te nemen van de bezoeken die respectievelijk 12, 9 en 6 maanden voor een ziekteopvlamming lagen. Deze genormeerde PROs werden vergeleken met de PROs ten tijde van een ziekteopvlamming, de PROs 3 maanden voor een opvlamming en de PROs 3-maandelijks na een opvlamming. Het bleek dat een ziekteopvlamming een negatief effect had op het algemene welzijn, ochtendstijfheid, functioneren, kwaliteit van

leven en vermoeidheid. Deze negatieve effecten waren vaak ook klinisch relevant en hielden tenminste 6 maanden aan. Alhoewel op groepsniveau de negatieve effecten van een ziekteopvlamming niet altijd groot lijken, kan dit voor een individuele patiënt alsnog van groot belang zijn.

Een belangrijke reden om medicatie af te bouwen is om kosten te besparen. In **hoofdstuk 9** onderzochten we daarom het afbouwen van medicatie vanuit maatschappelijk perspectief. We vergeleken de kosteneffectiviteit van de twee afbouwstrategieën in de TARA studie. Gezondheidszorgkosten waren lager in de groep patiënten die de TNF-inhibitor als eerste afbouwden, maar de patiënten in deze groep hadden daarnaast hogere maatschappelijke kosten als gevolg van productiviteitsverlies (meer ziekteverzuim en toename werkloosheid). Onderaan de streep waren dan ook de totale kosten vergelijkbaar tussen beide afbouwstrategieën. Daarnaast hebben we voor verschillende niveaus van betalingsbereidheid (willingness-to-pay, WTP) gekeken welke van de twee afbouwstrategieën het meest kosteneffectief is. De uitkomsten hiervan lagen zeer dicht bij elkaar. Het afbouwen van eerst de csDMARDs of eerst de TNF-inhibitor is daardoor even kosteneffectief en onafhankelijk van de WTP.

Hoofdstuk 10 bevat een algemene discussie van hoofdbevindingen van alle voorgaande hoofdstukken. De hoofdbevindingen worden besproken vanuit klinisch-, patiënt- en maatschappelijk perspectief. De nieuw verkregen inzichten en de mogelijke implicaties voor de dagelijkse praktijk worden hier besproken. Ook worden methodologische overwegingen bediscussieerd en mogelijke gevolgen daarvan op de resultaten. Als laatste worden aanbevelingen gedaan voor de reumatologische praktijk en worden er voorstellen gegeven voor toekomstig onderzoek. De aanbevelingen zijn dat afbouwen van medicatie kan worden gestart wanneer patiënten ten minste 6 maanden een DAS<2.4 hebben en eveneens geen gezwollen gewrichten. De volgorde voor afbouwen is niet relevant, vanwege invoegen meer bijwerkingen rapporteren van de csDMARDs, dat de meeste patiënten deels hun medicatie kunnen afbouwen, vooral in de groep die begon met afbouwen van csDMARDs, en de dalende biological prijzen, geef ik het advies om de csDMARD eerst af te bouwen. Daarnaast is mijn advies om gradueel af te bouwen en de patiënten goed te monitoren, ook met behulp van PROs.

ACKNOWLEDGEMENTS

I especially thank the participating patients in the TARA trial for their willingness to contribute to the study and for their cooperation. I thank all rheumatologists from the following participating centers: Erasmus MC, Rotterdam; Maasstad ziekenhuis, Rotterdam; Sint Fransiscus Gasthuis & Vlietland, Rotterdam & Schiedam; Amphia ziekenhuis, Breda; Reinier de Graaf ziekenhuis, Delft; Antonius ziekenhuis, Utrecht; Bravis ziekenhuis, Bergen op Zoom & Roosendaal; Groene hart ziekenhuis, Gouda; Albert Schweitzer ziekenhuis, Dordrecht; Haga ziekenhuis, Den Haag; and Beatrix ziekenhuis, Gorinchem. I also thank all study-nurses, laboratory personnel, co-investigators and others who were involved with the TARA study.

The TARA trial was supported by an unrestricted grant from ZonMW.

PHD PORTFOLIO

Name	Elise van Mulligen
Department	Rheumatology
Research School	Netherlands Institute for Health Sciences (NIHES)
PhD period	July 2017 – June 2019
Promotors	Prof. dr. J.M.W. Hazes
	Prof. dr. A.H.M. van der Helm – van Mil
	Prof. dr. A.E.A.M. Weel
Copromotor	dr. P.H.P. de Jong

PhD training	Year	Workload (ECTS)
General academic and research skills		
Basiscursus Regelgeving Klinisch Onderzoek (BROK)	2018	1.5
Workshop EndNote	2018	0.3
Workshop Systematic Literature Retrieval Pubmed	2018	0.3
Workshop Systematic Literature Retrieval in multiple databases	2018	0.3
Integrity in Scientific Research	2018	0.3
Introduction to Medical Writing	2020	2
In-depth statistical courses, NIHES		
Principles of Research in Medicine and Epidemiology	2018	0.7
Clinical Trials	2018	0.7
Methods of Public Health Research	2018	0.7
Fundamentals of Medical Decision Making	2018	0.7
Practice of Epidemiologic Analysis	2018	0.7
Health Economics	2018	0.7
Study Design	2018	4.3
Biostatistical Methods I: Basic principles	2018	5.7
Biostatistical Methods II: Classical Regression Models	2018	4.3
Intermediate course in R	2019	1.4
Quality of Life Measurement	2019	0.9
Repeated Measurements in Clinical Studies	2019	1.7
Logistic Regression	2019	1.4
Cohort studies	2019	0.7
Joint Models for Longitudinal and Survival data	2019	0.7
Clinical Translation of Epidemiology	2019	2.0

PhD training	Year	Workload (ECTS)
Clinical Epidemiology	2019	3.7
Principles in Causal Inference	2019	1.4
Advances analysis of Prognosis studies	2020	0.9
Missing Values in Clinical Research	2020	1.7
National conferences		
Nederlandse Vereniging voor Reumatologie (NVR) Najaarsdagen, Papendal [1 oral presentation, 1 poster presentation]	2018	2.0
Nederlandse Vereniging voor Reumatologie (NVR) Najaarsdagen, Papendal [1 oral presentation]	2019	1.0
International conferences		
European League Against Rheumatism (EULAR) Annual Meeting, Amsterdam, the Netherlands [2 oral presentations]	2018	2.0
European League Against Rheumatism (EULAR) Annual Meeting, Madrid, Spain [1 poster presentation]	2019	1.0
American College of Rheumatology (ACR) Annual Meeting, Atlanta, USA [1 oral presentation, 1 poster presentation]	2019	2.0
European League Against Rheumatism (EULAR) Annual Meeting, Frankfurt, Germany [1 oral presentation, 3 poster presentations]	2020	4.0
Seminars and workshops		
Department Journal Club Rheumatology	2018-2020	1.0
Department research meetings	2017-2020	1.0
Teaching		
Coaching of medical students	2018-2020	2.0
Other		
Study management TARA study	2017-2020	
Study management CSA study	2019-2020	
Co-organizer of PhD day Rheumatology	2019	
Grants		
Travel grant EULAR 2020, E-bursary	2020	

PUBLICATIONS

This thesis

van Mulligen E, Ahmed S, Weel A, Hazes JMW, van der Helm-Van Mil AHM, de Jong PHP. Factors that influence biological survival in rheumatoid arthritis: results of a real-world cohort from the Netherlands*. Submitted.*

van Mulligen E, de Jong PHP, Kuijper TM, van der Ven M, Appels C, Bijkerk C, Harbers JB, de Man Y, Molenaar THE, Tchetverikov I et al. Gradual tapering TNF inhibitors versus conventional synthetic DMARDs after achieving controlled disease in patients with rheumatoid arthritis: first-year results of the randomised controlled TARA study. *Ann Rheum Dis* 2019, 78(6):746-753.

van Mulligen E, Hazes JMW, Weel A, de Jong PHP. Response to: 'TARA Study: a new perspective on tapering drugs in RA' by Mishra et al. *Ann Rheum Dis* 2020;79:e80

Verstappen M*, **van Mulligen E***, de Jong PHP, van der Helm-Van Mil AHM. DMARD-free remission as novel treatment target in rheumatoid arthritis: A systematic literature review of achievability and sustainability. *RMD Open* 2020, 6(1):e001220.

van Mulligen E, Weel AE, Hazes JM, van der Helm-van Mil A, de Jong PHP. Tapering towards DMARD-free remission in established rheumatoid arthritis: 2-year results of the TARA trial. *Ann Rheum Dis* 2020;79:1174-1181.

van Mulligen E, de Jong PHP. Response to "Tapering towards DMARD-free remission in established rheumatoid arthritis: two year results of the TARA trial" by Haroon et al. *Ann Rheum Dis 2020.* Published Online First: 14 August 2020.

van Mulligen E, Weel A, Kuijper TM, Hazes JMW, van der Helm-van Mil AHM, de Jong PHP: The impact of a disease flare during tapering of DMARDs on the lives of rheumatoid arthritis patients. *Semin Arthritis Rheum. 2020;50(3):423-31*

van Mulligen E, Weel A, Kuijper TM, Denissen NHAM, Gerards AH, de Jager M, Lam WK, Hazes JMW, van der Helm-van Mil AHM, de Jong PHP*, Luime JJ*. Two-year cost-effectiveness between two gradual tapering strategies in Rheumatoid Arthritis: cost-utility analysis of the TARA trial. *Ann Rheum Dis 2020;79(12):1550-1556.*

Other

van Dijk CG, Nieuweboer FE, Pei JY, Xu YJ, Burgisser P, **van Mulligen E**, el Azzouzi H, Duncker DJ, Verhaar MC, Cheng C: The complex mural cell: pericyte function in health and disease. *Int J Cardiol* 2015, 190:75-89.

van Dijk CGM, Louzao-Martinez L, **van Mulligen E**, Boermans B, Demmers JAA, van den Bosch TPP, Goumans MJ, Duncker DJ, Verhaar MC, Cheng C: Extracellular Matrix Analysis of Human Renal Arteries in Both Quiescent and Active Vascular State. *Int J Mol Sci* 2020, 21(11).

ABOUT THE AUTHOR

Elise van Mulligen was born on the 5th of November 1991 in Utrecht, and grew up in Zeist. In 2010 she graduated Athenaeum at the Christelijk Lyceum Zeist and started with Biomedical Sciences at Utrecht University. After obtaining her bachelor's degree, she continued studying with the Master Regenerative Medicine & Technology at both Utrecht University and the Technical University of Eindhoven. After completion of two research projects and an extra minor about the fundamentals of business and economics, she obtained her Master degree in 2016.

After working elsewhere for one year she decided to start working at the department of Rheumatology of the Erasmus MC initially for half a year on the TARA project. After this first half year her contract was prolonged, and she started with her PhD under the supervision of Prof. dr. J.M.W. Hazes, Prof. dr. A.E.A.M. Weel, dr. P.H.P. de Jong, and later on also Prof. dr. A.H.M. van der Helm – van Mil.

During her PhD, she started the Research Master Health Sciences with the specialization Clinical Epidemiology from the NIHES (Netherlands Institute of Health Sciences), from which she graduated in August, 2020.

From July 2020 onwards, she started a postdoc under the supervision of Prof. dr. A.H.M. van der Helm – van Mil at the department of Rheumatology, Erasmus MC.

DANKWOORD

In dit laatste hoofdstuk is er gelukkig ruimte om iedereen te bedanken die heeft bijgedragen aan dit proefschrift.

Allereerst wil ik Pascal bedanken, bedankt voor je vertrouwen en daarnaast voor je altijd enthousiaste begeleiding en je geduld. Je was letterlijk altijd bereikbaar en mede dankzij jou heb ik het zo vlot kunnen voltooien, daarvoor heel veel dank! Hopelijk zijn er nog veel projecten die we samen kunnen gaan begeleiden in de toekomst.

Mieke en Angelique, ik vond onze overleggen altijd waardevol, jullie brachten mij altijd weer op nieuwe ideeën. Ik wil jullie daarvoor bedanken en ook voor het meedenken over mijn toekomst als onderzoeker.

Annette, ik wil je bedanken voor de kans die je mij hebt gegeven om mijn promotieonderzoek af te maken. Ik zie er naar uit om in de toekomst nog meer samen te gaan werken en ben heel erg blij dat ik als postdoc aan de slag kon.

Daarnaast wil ik ook alle TARA patiënten bedanken. Doordat jullie bereid waren om mee te doen aan het TARA onderzoek, heb ik er een mooi proefschrift over kunnen schrijven, dank daarvoor!

Martijn, Myrthe en Jolanda, bedankt voor al jullie werk aan de TARA studie, zonder jullie had ik hier nooit op kunnen promoveren. Martijn, ook nog veel dank voor het meedenken over de statistiek, fijn dat ik met vragen bij je terecht kon. Ook veel dank aan alle reumatologen en andere co-auteurs die hebben bijgedragen aan het TARA onderzoek.

Anke, bedankt voor zowel het verzamelen van data uit alle deelnemende centra voor de TARA studie, als ook voor je inzet voor de CSA studie. Ron, hoe vaak ik wel niet om een export heb gevraagd, toegang tot bepaalde mappen van de projectadministratie, het maken van nieuwe databases en vragenlijsten in GemsTracker: jij stond altijd klaar voor de nodige (IT) ondersteuning en altijd in voor een gezellig praatje, dank daarvoor! Joyce, bedankt voor het regelen van alle promotieoverleggen, de afdelingsuitjes en de coronanieuwsbrief.

Dan mijn onderzoekercollega's, heel erg bedankt, het was altijd heel gezellig op zowel de 6e als de 5e. Ik heb genoten van de vrijdagmiddelborrels, de PhD-dagen, het bezoek

aan Blijdorp, sushi eten, en alle andere uitjes. Ik ga geen volgorde aanbrengen, maar doe het op de welbekende manier: op lengte.

Jeroen, volgens jou komt alles goed. Het is heel fijn om jou als relativerende collega te hebben. Leuk om te zien dat we een beetje in hetzelfde tempo door onze promotie heen gaan. Bedankt dat je altijd de tijd neemt om mee te denken.

Luis, de EULAR in Madrid was super gezellig. Je weet me altijd te inspireren en je bent een goede observator. Fijn dat je me hebt aangemoedigd om te solliciteren bij EMEUNET, ik zie uit naar onze toekomstige podcasts.

Annelieke, jij bent de enige persoon in heel het Erasmus die de koffie lekker vindt, respect daarvoor. Ik vind je een hele fijne, gezellige collega en ben blij dat ik nog een poosje samen met jou postdoc kan zijn op de afdeling.

Hannah, ik bewonder jouw ambities en doorzettingsvermogen. Heel veel dank voor alle gezellige ganggesprekken. Ik hoop dat we in de toekomst ideeën kunnen blijven uitwisselen.

Margot, dank voor je nuchterheid en je altijd vrolijke humeur, het is gezellig om je bij de groep te hebben!

Nafise, jij hebt net een ander traject met ook je baan in het CBG, ik vind het knap hoe je dit combineert. Het was leuk om Iraans bij je te hebben gegeten!

Cleo, jij kwam als geroepen om samen met mij het CSA cohort nieuw leven in te blazen! Ik heb veel bewondering over jouw kennis en hoe jij alle artsen weet te motiveren. Ons wekelijkse kokoscappuccino moment is altijd een feestje en ik heb heel veel zin de komende twee jaar samen het CSA cohort verder te coördineren. Ik ben blij dat jij vandaag aan mijn zijde staat als paranimf.

Esther, jij bent de rots in de branding voor ons allemaal. Jij hebt zoveel kennis over zowel promoveren als over hoe je een studie opzet, maar ook hoe je met collega's om moet gaan. Het is heel fijn om jou als collega te hebben, bedankt voor al je steun!

Kim, heel erg bedankt dat ik altijd alles aan je kon vragen! Fijn dat je altijd mee wilde denken, voor elke Statavraag kon ik bij je terecht. Veel dank ook voor de gezellige borrels en de online zumbalessen.

Hilal en Rosaline, jullie kunnen sushi eten als de beste, verbazingwekkend hoeveel rondes jullie haalden. Hilal, bedankt dat ik altijd alles kon vragen, zeker in het begin was dat heel handig. En Rosaline, in het begin nog even worstelen met de hoogte van het bureau, maar daarna was het altijd gezellig om naast je op de kamer te zitten, bedankt!

Fazira, geen (teams)vergadering zonder snacks! Het is altijd gezellig om even bij te praten, ook al zit soms niet alles mee. Heel erg bedankt voor je humor en je gezelligheid.

Nathalie, we zijn tegelijk gestart (zonder bureau), bedankt voor de gezelligheid tijdens werk maar ook tijdens de NIHES cursussen en natuurlijk de pauzewandelingen naar de Kruidvat!

Verder nog alle andere collega's van de afdeling Reumatologie: bedankt! En dan nog veel dank aan mijn collega's uit Leiden: Marloes, Xanthe, Fenne, Yousra en Ellis. Leuk om ervaringen uit te wisselen en ook fijn dat ik op zowel de EULAR als de ACR bij jullie kon aanhaken.

Lieve vrienden en familie, ik wil jullie allemaal heel erg bedanken! Ik ben blij dat ik zoveel leuke en lieve mensen om mij heen heb en naast mijn werk doordeweeks heb ik aan jullie de nodige afleiding gehad. Dit in de vorm van barbecues, etentjes, weekendjes weg, feestjes, verjaardagen, tennistoernooien, of vakanties. Leonie, ik ben heel blij dat je vandaag als paranimf naast me wil staan, zeker omdat we al bevriend zijn sinds het prille begin van mijn wetenschappelijke carrière.

Lieve Benjamin, jij bent er altijd voor me en ik wil je ontzettend bedanken dat je me altijd hebt gesteund, van het vinden van een baan tot het afronden van mijn promotie. We zijn samen een goed team en ik zie heel erg uit naar onze bruiloft!