

**Prediction of response to therapy in rheumatoid arthritis**  
**Lost in validation**

Bart V.J. Cuppen

The studies presented in this thesis were conducted at the department of Rheumatology & Clinical Immunology, University Medical Center Utrecht and participating centers of the Society for Rheumatology Research Utrecht (SRU).

The research described in this thesis was financially supported by The Netherlands Association for Health Research and Development (ZonMW, project “Efficiency Studies: High cost medicines”), Foundation Friends of the UMCU, Pfizer, BMS, UCB, Roche, and Schering-Plough.

**Cover:** Rick Wuisman

**Lay-out:** Marieke Vianen en Bart Cuppen

**Printed by:** Ipskamp Printing

**ISBN:** 978-90-393-6745-2

© 2017, B.V.J. Cuppen, the Netherlands. All rights reserved. No part of this thesis may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system, without prior permission of the holder of the copyright.

Prediction of response to therapy in rheumatoid arthritis

Lost in validation

Predictie van respons op therapie in reumatoïde artritis: zoekgeraakt in validatie  
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van Doctor aan de Universiteit Utrecht,  
op gezag van de Rector Magnificus, prof. dr. G.J. van der Zwaan,  
ingevolge het besluit van het College voor Promoties  
in het openbaar te verdedigen op donderdag 23 maart 2017  
des ochtends te 10.30 uur

door

Bart Vincentius Joseph Cuppen  
geboren op 17 juni 1988 te Nijmegen

**PROMOTOREN:**

Prof. dr. F.P.J.G. Lafeber

Prof. dr. J.W.J. Bijlsma

**COPROMOTOREN:**

Dr. R.D.E. Fritsch-Stork

Dr. A.C.A. Marijnissen

## TABLE OF CONTENTS

1.	General introduction	1
2.	Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability	13
3.	Necessity of TNF-alpha inhibitor discontinuation in rheumatoid arthritis is predicted by smoking and number of previously used biological DMARDs	43
4.	Can baseline serum microRNAs predict response to TNF-alpha inhibitors in rheumatoid arthritis?	75
5.	RNA-sequencing to predict response to TNF- $\alpha$ inhibitors reveals possible mechanism for non-response in smokers	107
6.	CD4+ T cell mRNA evaluation by microarray and sequencing reveals no robust predictors of response to TNF-alpha inhibitors	155
7.	Proteomics to predict the response to TNF-alpha inhibitors in rheumatoid arthritis using a supervised cluster-analysis based protein score	185
8.	Exploring the inflammatory metabolomic profile to predict response to TNF- $\alpha$ inhibitors in rheumatoid arthritis	211
9.	Differences between serum polar lipid profiles of male and female rheumatoid arthritis patients in response to glucocorticoid treatment	257
10.	Polymorphisms in the multidrug-resistance 1 gene related to glucocorticoid response in rheumatoid arthritis treatment	279
11.	Summary of results and general discussion	293
12.	Nederlandse samenvatting	318
	Dankwoord	330
	Curriculum vitae	336
	List of publications	338



## **Chapter 1**

### **General introduction**

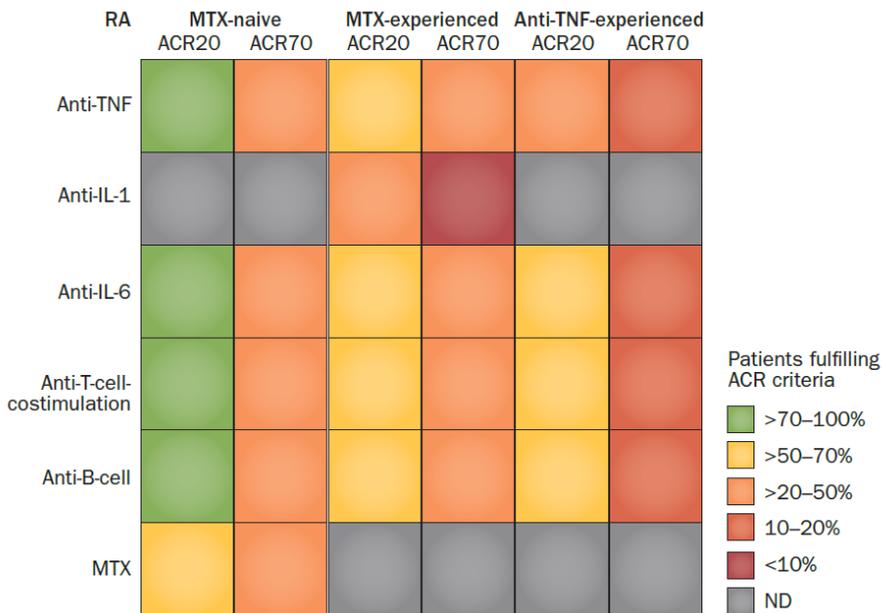


## **RHEUMATOID ARTHRITIS**

Rheumatoid Arthritis (RA) is a multifactorial chronic autoimmune disorder affecting 0.5-1.0% of the Western adult population [1-3]. In RA, overproduction of proinflammatory cytokines drives persistent synovial tissue inflammation and via proteases destruction of cartilage and subchondral bone [4]. Besides general complaints as fatigue, the more characteristic features of high disease activity include swelling, pain, and function loss of the affected joints with eventually structural damage [5]. The combination of a highly prevalent chronic disease with a high morbidity, makes RA a notable health issue in 21st century.

Although genetic and environmental risk factors for RA have been identified, the exact etiology of RA remains incompletely known [6, 7]. Likewise, many important players of the pathophysiological process have been identified, such as T-cells, B-cells, monocytes/macrophages and fibroblasts which execute their function via many different cytokines, yet these processes are very heterogeneous among patients [7, 8]. Still, the evolving understanding of this complex and extensive network has enabled the search for therapeutics with different targets, lowering disease activity by (partially) blocking diverse inflammatory pathways. These anti-inflammatory treatments, the so called Disease-modifying antirheumatic drugs (DMARDs), comprise a heterogeneous collection of agents and form the basis of RA therapy. Basically, DMARDs can be divided into two categories: the conventional synthetic DMARDs (csDMARDs) and the biological DMARDs (bDMARDs). The csDMARDs are small active-substance molecules which have been used for decades in the treatment of RA, and include drugs such as methotrexate (MTX), sulfasalazin (SSZ), leflunomide (LEF) and hydroxychloroquine (HCQ) [9]. The newer bDMARDs or biologicals are monoclonal antibodies that target and inhibit specific inflammatory pathways. From a clinical, scientific, and commercial perspective biologicals are one of the biggest successes of rational drug design, and have revolutionized the treatment options for RA patients since the first trial in 1993 [10-12]. Among biologicals, there are five registered agents inhibiting tumour necrosis factor alpha (TNF inhibitors (TNFi's)), these are: adalimumab (ADA), etanercept (ETN), infliximab (IFX), golimumab (GLM), and certolizumab pegol (CZP). Other biological agents, also called non-TNFi's, target IL-6 receptor (tocilizumab (TCZ)), T-cell co-stimulation CD80/CD86 (abatacept (ABA)), B lymphocyte antigen CD20 (rituximab

(RTX)) or IL-1 receptor (anakinra (ANA)). The effective control of disease activity by biologicals is clearly visible when the clinical response rates to MTX and the biologicals are observed (see **Figure 1**)[13]. Of note, newer agents for RA have been introduced that necessitated DMARD subdivision into more categories, such as targeted synthetic DMARDs (e.g. tofacitinib; an inhibitor of JAK enzymes, interfering with transduction of cellular pathways that are normally activated by proinflammatory cytokines) or biosimilar DMARDs (e.g. bs-infliximab; sharing most but not all features of the parent compound)[9]. As current treatments only manage to suppress RA disease activity with variable success, novel molecules are constantly under investigation.



**Figure 1:** Clinical responses to methotrexate (MTX) and several biologicals, as subdivided for MTX-naive (usually early RA), MTX-experienced and TNF-alpha inhibitor experienced (established RA). The response to bDMARDs is superior to MTX, whereas the effectiveness of the bDMARDs (except for anti-IL-1) is equal for patients with a comparable drug-experience. The American College for Rheumatology (ACR) response criteria indicate an improvement of  $\geq 20\%$  and  $\geq 70\%$  for ACR20 and ACR70, respectively. ND, not done; RA, rheumatoid arthritis; TNF = tumour necrosis factor. Figure by Smolen et al. [13].

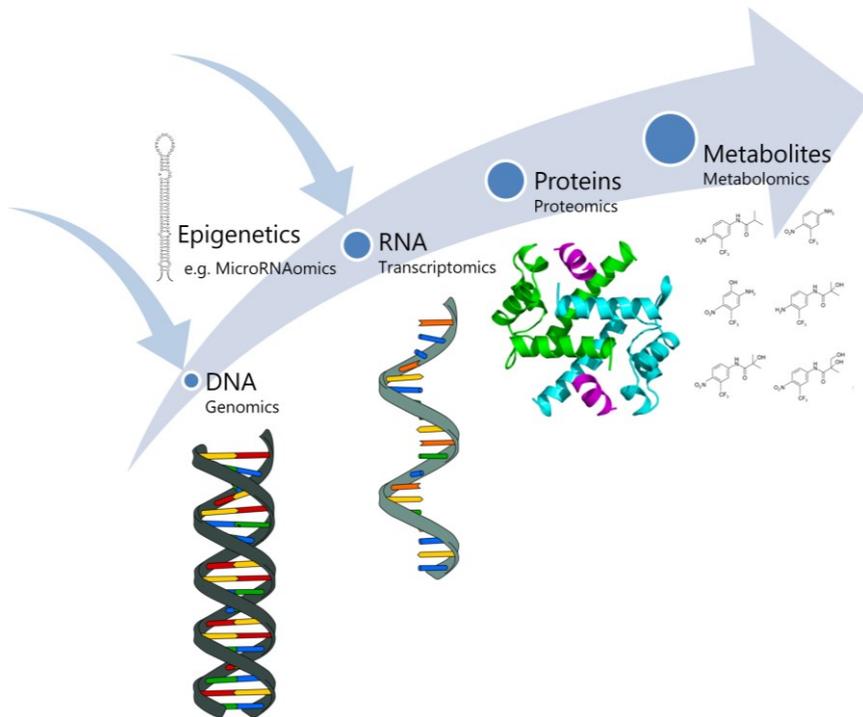
## **TRIAL-AND-ERROR APPROACH**

When a patient is diagnosed with RA, the initial DMARD treatment is usually MTX, preferably in combination with low-dose glucocorticoids. If this treatment and/or other (combinations of) csDMARDs fail, and a patient has a high risk on rapid radiographic progression (i.e. high disease activity)[9, 14], current strategy is to prescribe a biologic [9]. After commencing a new biological the clinical response may take up to several months, still approximately 30-40% of the RA patients will clinically not sufficiently respond to their first biological (**Figure 1** and [15, 16]). Therefore, if the clinical improvement is limited after 3 months of therapy, non-responders are encouraged to switch to another biological[7] and the process of awaiting clinical response is repeated. This trial-and-error approach is very inefficient, with patients suffering from uncontrolled disease with the potential of irreversible joint damage and the risk of harmful side effects. In addition but not less important, as treatment with biologicals is costly [17] and non-response eventually results in an increased morbidity, the trial-and-error approach unavoidably incorporates expenditures of valuable resources. Hence there is a need for a more refined treatment strategy, ideally wherein response to treatment can be predicted before the start of treatment and guide the optimal choice for each patient. Such a prediction may be related to (a combination of) patient characteristics, a different phenotype or phase of the disease. When, for example a patient is eligible for a biological and the response to a TNFi is predicted, a patient with a high chance of response will be encouraged to start treatment, whereas other treatment options (also non-biological DMARDs) might be considered in a patient with a likely non-response. Personalised medicine in this sense is already well-established in the field of oncology, such as measurement of the Her2 receptor expression in breast cancer tissue to predict response to trastuzumab [18, 19]. These successes have motivated researchers to start a quest for predictors of response in RA as well, aiming at the development of a more personalised treatment approach to reduce disease burden, diminish treatment risks, and consequently reduce costs to maintain resources for those who benefit best.

## PREDICTION OF RESPONSE

The search for predictors of response started with exploring readily available clinical parameters in routine care, of which many have been explored to date [20]. Basically, the only success was the finding that absence of the serologic autoantibodies rheumatoid factor and anti citrillunated protein antibodies can, very marginally, predict a decreased chance of response to RTX [21]. However, no recommendations for a tailored treatment approach according to serologic autoantibody status are embedded in the official guidelines [9, 14], likely because the differences in response according to serological status are too small and RTX is still considered an effective biological in all RA patients and a good first choice after MTX failure [7]. Of note, but more a form of bias than a true predictor, the disease activity at baseline strongly influences the chances of response because the change in disease activity over time is usually incorporated in the response definition. This is not always fair: when for example a patient is eligible for biological treatment but in the week prior to initiation the disease activity has decreased already because glucocorticoids were recently added, chances of response will also arbitrarily decrease. In addition, patients with a very high baseline disease activity, have a substantial higher chance of lower values at any time after that due to random fluctuations, a concept known as regression to the mean. The disease activity is therefore rather a factor to take into account in prediction analysis than a useful and relevant predictor for clinical practice.

Due to the limited additive success achieved by clinical parameters to discriminate between responders and non-responders to biologicals, the focus shifted to biochemical markers more closely related to RA pathogenesis, arising from the cascade DNA, RNA, epigenetics, proteins, and metabolites (see **Figure 2**). Since then, the field of genetics, transcriptomics, microRNAomics, proteomics, and metabolomics are being explored by scientists in the wish and hope of identifying those factors that differ at baseline between responders and non-responders. However, the still unresolved heterogeneity in RA disease processes and in clinical response to therapy, together with the complexity of and interaction between the different omics fields that are still rapidly evolving, makes prediction of therapeutic response one of the major challenges in RA.

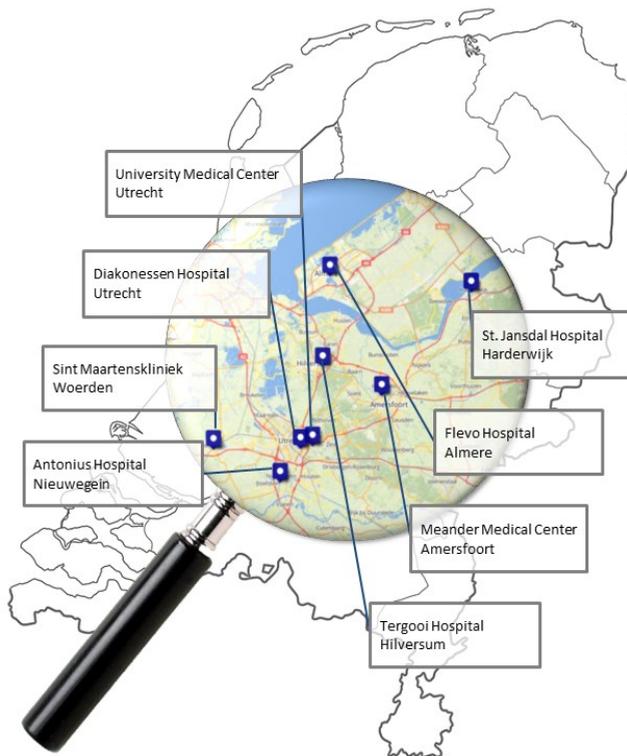


**Figure 2:** The fields of ‘omics’. Biomarkers can be investigated from the complete downstream DNA cascade, which involves biological molecules that (eventually) translate into structure, function, and dynamics of the human body. Each step in this cascade has its field of ‘omics’, which denotes the collective of the biological molecules involved. These omics fields have been targeted for many diseases, including RA. The illustrations of biomarkers used in this figure are all derived from Wikipedia.org.

## AIMS AND OUTLINES OF THIS THESIS

Driven by the unmet desire from clinical practice, this thesis aims to identify the most valuable viz. clinically relevant (combination of) predictors of response to bDMARD therapy in RA, by careful evaluation of existing literature and on original study data in combination with high quality experiments on patient material. These data and materials were mainly derived from the BiOCURA cohort: Biologicals and Outcomes, Compared and predicted, in Utrecht region, in Rheumatoid Arthritis. In line with the aim of this thesis, this cohort was started as a means to find a prediction rule for response to biological therapy. The BiOCURA is a multi-center cohort in collaboration with eight hospitals within

the Society for Rheumatology Research Utrecht (SRU) (see **Figure 3**). In a time span of six years (2009-2015), over 400 RA patients commencing biological treatment were included. Of all patients, clinical data and specimens (blood and urine) were collected before treatment initiation, and patients were followed for one year after start of treatment in order to assess the clinical response (efficacy) and to monitor drug safety.



**Figure 3:** SRU hospitals and their locations in the Utrecht region. The SRU is a collaboration of hospitals that joined efforts to include patients in the BiOCURA cohort. This has resulted in more than 400 inclusions in 6 years.

### *What is known about predictive markers of response to treatment in RA?*

To answer this question we conducted a systematic review. We evaluated the existing literature addressing prediction of response to bDMARDs in RA, by performing a

systematic review as described in **chapter two**. In an effort to address the usefulness of all published predictors, we calculated with the derived original data from these studies to obtain similar - and thus comparable - predictive abilities.

#### *Who will (dis)continue TNFi treatment?*

As biological treatment in BIOCURA was frequently discontinued due to inefficacy and/or side effects, treatment success as a whole, continuation versus discontinuation, was explored in **chapter three**. Clinical characteristics of patients who discontinued TNFi were studied and compared with previous studies that also evaluated predictors for discontinuation of TNFi treatment.

#### *Can we find better biomarkers:*

##### *1) By using microRNAs?*

MicroRNA are small non-coding RNA molecules that function in RNA silencing and post-transcriptional regulation of gene expression. In **chapter four**, candidate predictors for the therapeutic response to TNFi treatment were selected from a large panel (>750 miRNAs). After identification, a second separate cohort was used in order to validate the findings.

##### *2) By using messenger RNA (mRNA)?*

mRNAs, which convey genetic information from DNA to be translated into proteins, derived from white blood cells were evaluated as a marker for prediction of response to TNFi treatment in **chapter five**. After determination of potential targets, measurement of these expressed genes in a separate cohort was used to validate the identified predictors. Since Th-cells are a subtype of white blood cells and important players in RA pathogenesis, the quest for mRNA of Th-cells as possible source for predictors of response to TNFi treatment is described in **chapter six**. In three partially different patient cohorts, the differential expression of gene transcripts was determined by two different profiling techniques: microarray and RNA-sequencing. Overlap among the experiments was used to cross-validate any interesting target.

### *3) By using proteomics?*

Next in **chapter seven**, a proteomic approach was employed to predict response to TNFi therapy. The proteasome was first investigated for the best combination of biomarkers to develop a prediction rule, which was then tested for its additional value over prediction using clinical parameters alone. Subsequently, the prediction rule was validated in a separate cohort to investigate its replicability.

### *4) By using metabolomics?*

Lastly, metabolomics in combination with clinical parameters was tested for its ability to predict response to TNFi therapy, as described in **chapter eight**. Since clinical parameters influence the measured biomarkers to a certain extent, the influence of clinical parameters on metabolomics was further investigated in **chapter nine**, in particular for the role of concomitant use of glucocorticoids (GC) and the levels of lipids.

### *5) By using genomics (in GC response)?*

Additionally in **chapter ten**, gene polymorphisms in the multidrug resistance-1 gene were studied in the context of response to intravenous and oral glucocorticoids.

### *How far are we with prediction?*

Finally the contribution of this work towards a more personalised treatment approach in RA is discussed in **chapter eleven**. The main question to address is: are we anywhere closer in finding predictors and can we fully succeed to predict response to (biological) therapy in RA? If so, these predictors might also further the understanding of the mechanisms that drive disease activity and potentially reveal new treatment targets for RA. Finally, ideas to increase the quality and reliability of future studies on this topic are deliberated on.

## REFERENCES

1. Wolfe AM, Kellgren JH, Masi AT. The epidemiology of rheumatoid arthritis: a review. II. Incidence and diagnostic criteria. *Bull Rheum Dis* 1968;19 3: 524-9.
2. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res* 2002;4 Suppl 3: S265-S72.
3. Gabriel SE, Michaud K. Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. *Arthritis Res Ther* 2009;11 3: 229.
4. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. *Lancet* 2010;376 9746: 1094-108.
5. Bijlsma JWJ, van Laar JM. *Leerboek reumatologie en klinische immunologie: Bohn Stafleu van Loghum*, 2013.
6. Jutley G, Raza K, Buckley CD. New pathogenic insights into rheumatoid arthritis. *Curr Opin Rheumatol* 2015;27 3: 249-55.
7. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet* 2016.
8. Goronzy JJ, Weyand CM. Developments in the scientific understanding of rheumatoid arthritis. *Arthritis Res Ther* 2009;11 5: 249.
9. Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73 3: 492-509.
10. Elliott MJ, Maini RN, Feldmann M, Long-Fox A, Charles P, Katsikis P et al. Treatment of rheumatoid arthritis with chimeric monoclonal antibodies to tumor necrosis factor alpha. *Arthritis Rheum* 1993;36 12: 1681-90.
11. [editorial]. Biologic drugs set to top 2012 sales. *Nat Med* 2012;18 5: 636.
12. Tansey MG, Szymkowski DE. The TNF superfamily in 2009: new pathways, new indications, and new drugs. *Drug Discov Today* 2009;14 23-24: 1082-8.
13. Smolen JS, Aletaha D. Rheumatoid arthritis therapy reappraisal: strategies, opportunities and challenges. *Nat Rev Rheumatol* 2015;11 5: 276-89.
14. Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol* 2016;68 1: 1-26.
15. Keystone EC, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum* 2004;50 5: 1400-11.
16. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med* 1999;340 4: 253-9.
17. Kievit W, Adang EM, Franssen J, Kuper HH, van de Laar MA, Jansen TL et al. The effectiveness and medication costs of three anti-tumour necrosis factor alpha agents in the treatment of rheumatoid arthritis from prospective clinical practice data. *Ann Rheum Dis* 2008;67 9: 1229-34.
18. Mass RD, Press MF, Anderson S, Cobleigh MA, Vogel CL, Dybdal N et al. Evaluation of clinical outcomes according to HER2 detection by fluorescence in situ hybridization in women with metastatic breast cancer treated with trastuzumab. *Clin Breast Cancer* 2005;6 3: 240-6.
19. Ross JS, Slodkowska EA, Symmans WF, Puzstai L, Ravdin PM, Hortobagyi GN. The HER-2 receptor and breast cancer: ten years of targeted anti-HER-2 therapy and personalized medicine. *Oncologist* 2009;14 4: 320-68.
20. Daïen CI, Morel J. Predictive factors of response to biological disease modifying antirheumatic drugs: towards personalized medicine. *Mediators Inflamm* 2014;2014: 386148.
21. Isaacs JD, Cohen SB, Emery P, Tak PP, Wang J, Lei G et al. Effect of baseline rheumatoid factor and anticitrullinated peptide antibody serotype on rituximab clinical response: a meta-analysis. *Ann Rheum Dis* 2013;72 3: 329-36.



## Chapter 2

### **Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability**

Bart V.J. Cuppen<sup>#</sup>, Paco M.J. Welsing, Jan J. Sprengers, Johannes W.J. Bijlsma, Anne C.A. Marijnissen, Jacob M. van Laar, Floris P.J.G. Lafeber, Sandhya C. Nair<sup>#</sup>

# These authors equally contributed

## **ABSTRACT**

### **Objective**

To review studies that address prediction of response to biological treatment in rheumatoid arthritis (RA) and to explore the clinical utility of the studied (bio)markers. The present data suggests that local subclinical inflammation occurs in ACPA positive

### **Methods**

A search for relevant articles was performed in PUBMED, EMBASE and COCHRANE databases. Studies which presented predictive values or in which these could be calculated were selected. The added value was determined by the added value on prior probability for each (bio)marker. Only an increase/decrease in chance of response of  $\geq 15\%$  was considered clinically relevant enough, whereas in oncology values  $>25\%$  are common.

### **Results**

Out of the 57 eligible studies, 14 (bio)markers were studied in more than one cohort and an overview of the added predictive value of each marker is presented. Of the replicated predictors, none consistently showed an increase/decrease in probability of response of  $\geq 15\%$ . However, positivity of RF and ACPA in case of RTX and presence of TNF- $\alpha$  promotor -308 GG genotype for TNFi therapy were consistently predictive yet low in added predictive value. Besides these, 65(bio)markers studied once, showed remarkably high (but not validated) predictive values.

### **Conclusion**

We were not able to address clinically useful baseline (bio)markers, for use in individually tailored treatment. Some predictors are consistently predictive yet low in added predictive value, while several others are promising but await replication. The challenge now is to design studies to validate all explored and promising findings individually and in combination, to make these (bio)markers relevant to clinical practice.

## INTRODUCTION

Rheumatoid arthritis (RA) is a common multi-factorial autoimmune disorder affecting one in 100-200 persons in the Western world [1]. The disease frequently results in reduced mobility, since it predominantly affects patients' joints. Therapy consists of immunosuppressive drugs in combination with pain management, life-style changes and physiotherapy [2]. Therapeutic interventions have advanced tremendously with the introduction of biologicals targeting pivotal mediators in the inflammatory process, e.g. by inhibiting tumour necrosis factor alpha (TNF inhibitors, TNFis), interleukin-1 (IL-1, anakinra), interleukin-6 (IL-6, tocilizumab (TCZ)), T-cell co-stimulation CD80/CD86, (abatacept) or B-lymphocyte antigen (CD20, rituximab (RTX)). As a consequence, disease burden has decreased due to more effective control of disease activity with a substantial reduction in joint damage. Biological treatments have a specific mechanism of action and patients possibly will respond best to the biological that acts on the target that predominantly mediates the current inflammatory process in the individual patient. This implies an optimal choice for a biological treatment for the individual patient and thus a need for personal-, viz. patient tailored treatment. At present, of all patients receiving biologicals 30-40% do not respond to the prescribed biological [3, 4]. In addition, these treatments are associated with possible (serious) adverse effects (e.g. opportunistic infections and toxicity risk) [5]. Ineffective treatment with these biologicals as well as adverse effects might deteriorate the patients' well-being and thus undermine the primary goal of therapy. Current strategy in clinical practice is to prescribe a biological when initial treatment with methotrexate (MTX), often in combination with low-dose glucocorticoids, fails. When the decrease in disease activity after three to six months is insufficient for the first biological (usually a TNFi), the patient switches to another biological and the process to await clinical response is repeated. This trial and error approach accompanies inefficiency, with patients suffering from uncontrolled disease and the potential of irreversible joint damage. In addition, treatment with biologicals is (very) costly [6], which results in expenditure of valuable resources. Hence, there is a need for a more refined strategy, wherein response to specific treatments can be predicted before the start of treatment and guide the optimal choice for each particular patient. Personalized treatment could reduce disease burden, diminish treatment risks and

consequently reduce healthcare costs from monitoring, complications and morbidity. Prediction of response in oncological diseases is already well-established, such as Her2 receptor expression and response to trastuzumab in breast cancer [7]. Approximately 15-20% of the breast cancer patients expresses this receptor, which predicts a considerable higher probability of response to trastuzumab [8].

Several studies have used different strategies to identify (bio)markers in RA patients that could predict treatment outcome and therefore could guide physicians in their decision to prescribe the most suitable biological. Scientists employ the quickly developing fields of genetics, proteomics, and transcriptomics in combination with clinical variables to identify those factors that differ at baseline between responders and non-responders.

In this systematic review, prediction studies are compiled in order to provide an overview of the proposed predictors currently investigated for several biologicals. The aim is to explore the added value in clinical applicability of these (bio)markers in predicting response to biological treatment in RA.

## **METHODS**

### **Search strategy**

In this review three researchers, SN, BC and JS, independently defined their search strategy for PUBMED. These strategies were compiled and thereafter controlled for redundancy to synthesize one final search string (see Supplementary Text 1). The core components of this string consisted of the definition of the patient group (RA-patients starting treatment with a biological), measure of interest (prediction of response) and outcome parameters (response criteria). Subsequently, articles were obtained from multiple databases (i.e. PUBMED/MEDLINE, EMBASE, and COCHRANE) until February 2014. The retrieved articles were screened by at least two authors (SN, BC, JS) for title and abstract. Only articles written in English, concerning human studies with RA patients, treated with any biological and predicting response defined by 28 joint count Disease Activity Score (DAS28), European League Against Rheumatism (EULAR) or American College of Rheumatology (ACR) response criteria were included for further full-text analysis. To prevent the exclusion of valuable articles, articles were included for full-text

analysis if there was uncertainty about meeting these inclusion criteria (e.g. the abstract did not specify the response criteria used). Articles were excluded during the full text analysis when the follow-up time for response was shorter than 12 weeks, the (bio)marker/variable/test was determined more than 4 weeks after initiation of the treatment (i.e. not measured at baseline), the study only observed changes of a marker over time under influence of the biological (i.e. no result regarding prediction), or if the article concerned a review. To check for articles missed by the search strategy cross reference checking was done to check for missed papers, which were tested using the before mentioned in- and exclusion criteria.

### **Data extraction**

Information on study design, baseline patient/disease characteristics and results on proposed predictors was extracted from individual studies. Quantitative information on the number of patients with a positive and negative test value (i.e. biomarker above/below the cut-off value) and of these the number of patients with a positive and a negative outcome (i.e. response/non-response) were extracted from tables with unadjusted parameters. Although in some studies multivariable models are (also) used to describe the effect of the markers on response, these models were not included in this study because different sets of biomarkers and clinical correction factors are used and solely no quantitative information on test positive and -negative patients can be extracted for an individual predictor. Consequently, comparing predictors between studies is both biased and impossible. Multibiomarker models without the combination of clinical correction factors were selected to be presented separately from the individual predictors.

For the EULAR response criteria[9] moderate and good responders were pooled and contrasted to non-responders and for the ACR response criteria[10], ACR-20 was extracted if possible. These cut-offs were preferred for comparison since a smallest response might result in a striking relief in disease burden as experienced by the patient. Studies that reported an absolute change in DAS28-score of 1.2 were included, because this includes both good and moderate responders. Overall, articles which predict EULAR good response or ACR-70 only were excluded, first because the predictive values could not be compared

to articles which predict any response at all and second, because in clinical practice quite some patients only reach moderate responses.

### **Predictor selection**

We divided the predictors in either one of the following two groups:

- Markers/variables studied multiple times (MSMs) - These are the extracted predictors reported in 2 or more of the included studies (Table 1).
- Markers/variables studied once (MSO) - These are all extracted predictors that are reported only in a single included study (Table 2).

In several articles, besides the primary investigated predictors, additional information on (clinical) factors with potential predictive ability was available. In these cases, quantitative information on these extra variables was extracted, provided they were addressed in other studies (MSMs).

We were interested in comparing (and quantitatively summarize) the results of different studies. As a consequence, when studies used multiple cut-off values for their predictor, we used the value that was most comparable to the other studies using that predictor. For example, when a gene has a C/G polymorphism and carriage of the CC genotype predicts response or non-response, the C homozygous patients are compared to all others, including patients with the CG and GG genotype. A risk of bias assessment was carried out for all included studies using the Quality In Prognosis Studies (QUIPS) tool[11]. All included studies were assessed by two authors (SN, BC, JS), after which results were compared and differences discussed until agreement was reached.

### **Statistical analysis**

The prior probability of response (i.e. the chance of an individual at baseline to obtain a response irrespective of the test values in a specific study population) was calculated by dividing the total number of responders by the total number of patients that entered the study. If it was not mentioned in the article what is test positive (or predicting response) and test negative (or prediction non-response), the calculated likelihood ratio (LR) was used to determine the direction of the effect. The positive predictive value (PPV, i.e. the proportion of test positives which are responders (true positives) among all test positives)

and negative predictive value (NPV, i.e. the proportion of test negatives which are non-responders (true negatives) among all test negatives) of each (bio)marker was then calculated. Contrary to sensitivity and specificity, the PPV and NPV are highly dependent on the (prior) probability of response in the patient population in which they are calculated. This means that the PPV and NPV calculated in two separate studies can be different when the chances of response in the two study populations are not comparable, even when the predictive ability of the predictors is the same in the two studies. Therefore, a weighted (on study size) pooled prior probability was calculated for studies based on class of biological and response criterion used. This technique has been used before to enhance comparability for heterogeneous studies [12]. Because a pooled prior probability was used, the posterior probability after a positive (=PPV) and negative test (1-NPV) were calculated again by applying the LR<sub>s</sub> of the (bio)markers on the pooled prior probability. The risk difference between the pooled prior probability and posterior probability was calculated to show the added predictive value (APV) of the test, after test positive and test negative (APV+ and APV- respectively). The APV+ and APV- show the increase and decrease in absolute chance of response when the test-outcome (high or low value of the (bio)marker) is known.

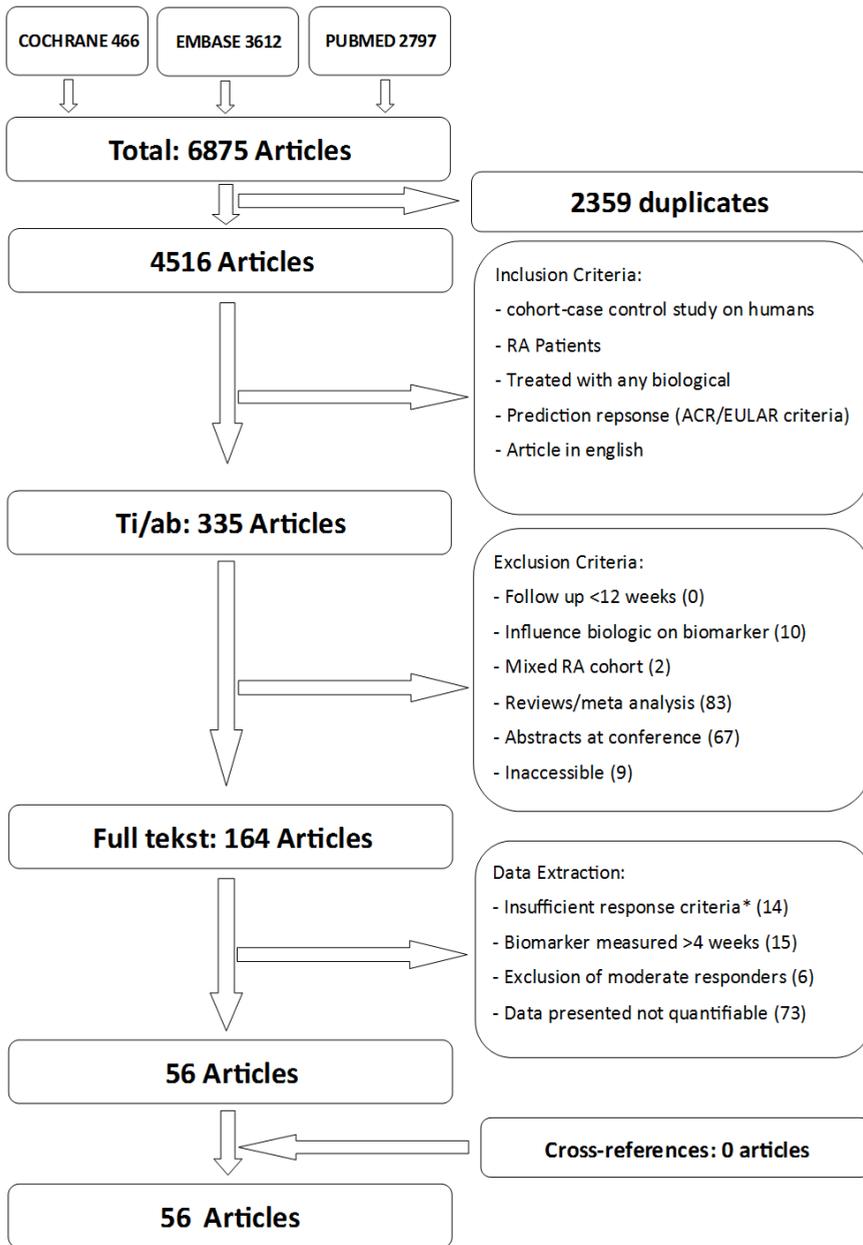
To determine the clinical relevance for MSMs and MSO, a cut-off of  $\geq 15\%$  for APV+ and APV- (i.e. a  $\geq 15\%$  higher probability of response after a positive test, or a  $\geq 15\%$  lower probability of response after a negative test) was chosen. This is not conservative, whereas in the case of trastuzumab and breast cancer, risk differences in the order of  $>25\%$  are found [13, 14]. In addition, besides the cut-off of  $\geq 15\%$  the clinical relevance for MSMs was also based on the consistency of the (bio)marker in predicting the 'same direction' over the different studies. Statistical analyses were performed in SPSS 21.0.0.0 (SPSS Inc., Chicago, Illinois, USA).

## **RESULTS**

### **Search**

The final search from the three databases yielded 4516 unique articles of which 164 were considered relevant based on title and abstract (see Figure 1). After full-text screening,

107 articles were excluded because response criteria did not meet the preselected criteria (n=14), markers were measured after 4 weeks of therapy initiation (n=15), moderate responders were excluded (n=6) or data was not quantifiable (e.g. no quantitative information or no cut-off value used) (n=72). Cross-reference checking did not result in extra articles for data-extraction. From the remaining studies, data on 79 (bio)markers (MSMs and MSO) [15-64] and 8 multivariable biomarker models [65-71] was extracted. The risk of bias was assessed using the QUIPS (see Supplementary Table 1).



\* response criteria presented does not meet our preselected response criteria

**Figure 1:** flowchart

### **Markers/variables studied multiple times (MSMs)**

Fourteen predictors were analysed in multiple studies: erosive disease, concomitant MTX use, smoking, rheumatoid factor (RF), three immunoglobulin classes of RF, anti-citrullinated protein antibody (ACPA), anti-nuclear antibody (ANA), Epstein-Barr virus (EBV)-status, immunoglobulin G (IgG) levels, Fc-gamma receptor IIIA (FCGR3A) 158V/F-, IL-6 promotor 174G/C- and TNF- $\alpha$  promotor 308G/A polymorphisms (Table 1) . These predictors have been tested in studies on different biologicals (RTX, TCZ and several TNFis) and response was measured using different response criteria (ACR20, ACR 50, EULAR and a decrease in DAS28-score of 1.2). As described before, test positivity was defined based on the outcome and should increase the chance of response. Prior probabilities were pooled for studies with similar response criteria and class of biological and PPV and 1-NPV were calculated, as well as APV+ and APV-

**Table 1:** Overview of predictors studied multiple times (MSMs). Studies could be present multiple times depending on number of MSMs that were presented. The prior probability pooled (i.e. the chance of an individual at baseline to obtain a response irrespective of the predictor) was based on actual observed response data, extracted from studies that could be pooled based on same treatment and same response criterion used.

Author (et al.)	Total no.	Prior. prob. per study, %	Response criteria	Class of biological	Predictors	Prior prob. pooled, %	Post. prob +, %	Post. prob -, %	APV+, %	APV-, %
Daïen [25]	63	69.8	EULAR	RTX	Erosive disease +	68.5	71.8	53.0	3.3	-15.5
Sellam [59]	208	71.6	EULAR	RTX	Erosive disease +	68.5	70.3	54.6	1.8	-13.8
Bos et al. [17]	188	77.1	EULAR	TNFi	Erosive disease +	71.2	72.6	65.6	1.4	-5.6
Braun-Moscovici [18]	30	80.0	$\Delta$ DAS28 1.2	TNFi	Erosive disease +	71.2	71.2	71.2	0.0	0.0
Klaasen [39]	101	68.3	$\Delta$ DAS28 1.2	TNFi	Erosive disease +	71.2	72.9	65.7	1.7	-5.6
Wijbrandts [63]	103	68.0	$\Delta$ DAS28 1.2	TNFi	Erosive disease +	71.2	75.2	58.0	3.9	-13.3
Gottenberg [30]	558	59.1	EULAR	ABA	MTX use -	59.1	60.5	58.7	1.4	-0.4
Couderc [23]	64	71.9	EULAR	RTX	MTX use +	68.5	71.8	65.2	3.4	-3.3
Soliman [58]	627	60.1	EULAR	RTX	MTX use +	68.5	69.3	67.6	0.8	-0.9
Abhishek [15]	395	89.4	EULAR	TNFi	MTX use -	71.2	79.5	67.0	8.3	-4.2
Hyrich [33]	3,223	67.8	EULAR	TNFi	MTX use -	71.2	73.7	67.6	2.4	-3.6
Bos [17]	188	77.1	EULAR	TNFi	MTX use +	71.2	77.9	53.7	6.6	-17.5
Braun-Moscovici [18]	30	80.0	$\Delta$ DAS28 1.2	TNFi	MTX use +	71.2	88.9	57.7	17.7	-13.6
Pers[50]	165	86.1	EULAR	TCZ	MTX use +	86.1	88.6	83.7	2.5	-2.3
Abhishek [15]	395	89.4	EULAR	TNFi	Not smoking	71.2	75.0	56.8	3.7	-14.4
Hyrich [33]	3,223	67.8	EULAR	TNFi	Not smoking	71.2	71.6	69.7	0.4	-1.5
Pers [50]	165	86.1	EULAR	TCZ	Currently smoking	86.1	88.9	82.7	2.8	-3.4
Gottenberg [30]	558	59.1	EULAR	ABA	RF +	59.1	62.2	51.1	3.1	-8.0
Quartuccio [53]	110	66.4	ACR 50	RTX	RF +	66.4	75.3	28.6	8.9	-37.8
Couderc [23]	64	71.9	EULAR	RTX	RF +	68.5	76.3	53.1	7.8	-15.4
Daïen [25]	63	69.8	EULAR	RTX	RF +	68.5	72.7	57.3	4.2	-11.2
Fabris [27]	30	80.0	EULAR	RTX	RF +	68.5	74.0	44.9	5.6	-23.6
Soliman [58]	627	60.1	EULAR	RTX	RF +	68.5	70.4	64.4	1.9	-4.1
Buch [19]	207	51.7	ACR 20	TNFi	RF +	58.6	69.2	21.9	10.6	-36.7
Stuhlmüller [61]	27	55.6	ACR 20	TNFi	RF +	58.6	64.8	36.1	6.2	-22.4
Bobbio-Pallavicini [16]	126	65.9	EULAR	TNFi	RF -	71.2	72.0	71.1	0.7	-0.1
Chen [21]	48	75.0	EULAR	TNFi	RF -	71.2	83.2	69.2	12.0	-2.0

<b>Abhishek [15]</b>	395	89.4	EULAR	TNFi	RF +	71.2	78.1	64.9	6.9	-6.3
<b>Hyrich [33]</b>	3,223	67.8	EULAR	TNFi	RF +	71.2	71.6	70.2	0.4	-1.0
<b>Morales-lara [46]</b>	77	57.1	EULAR	TNFi	RF +	71.2	84.0	67.7	12.8	-3.6
<b>Lequerre [42]</b>	33	48.5	$\Delta$ DAS28 1.2	TNFi	RF +	71.2	74.2	63.7	2.9	-7.5
<b>Wijbrandts [63]</b>	103	68.0	$\Delta$ DAS28 1.2	TNFi	RF +	71.2	76.9	53.9	5.7	-17.4
<b>Pers [50]</b>	165	86.1	EULAR	TCZ	RF -	86.1	92.2	83.3	6.1	-2.7
<b>Fabris [27]</b>	30	80.0	EULAR	RTX	IgA RF +	68.5	75.5	55.9	7.0	-12.6
<b>Bobbio-Pallavicini [16]</b>	126	65.9	EULAR	TNFi	IgA RF +	71.2	71.3	71.2	0.0	-0.1
<b>Klaasen [39]</b>	101	68.3	$\Delta$ DAS28 1.2	TNFi	IgA RF +	71.2	78.2	67.8	7.0	-3.4
<b>Bobbio-Pallavicini [16]</b>	126	65.9	EULAR	TNFi	IgG RF -	71.2	77.5	67.2	6.3	-4.0
<b>Klaasen [39]</b>	101	68.3	$\Delta$ DAS28 1.2	TNFi	IgG RF +	71.2	85.2	66.8	13.9	-4.5
<b>Bobbio-Pallavicini [16]</b>	126	65.9	EULAR	TNFi	IgM RF -	71.2	76.7	69.8	5.5	-1.4
<b>Bos [17]</b>	188	77.1	EULAR	TNFi	IgM RF +	71.2	75.0	66.3	3.8	-4.9
<b>Klaasen [39]</b>	101	68.3	$\Delta$ DAS28 1.2	TNFi	IgM RF +	71.2	80.1	56.1	8.8	-15.2
<b>Gottenberg [30]</b>	558	59.1	EULAR	ABA	ACPA +	59.1	63.9	48.0	4.7	-11.2
<b>Quartuccio [53]</b>	110	66.4	ACR 50	RTX	ACPA +	66.4	71.8	48.0	5.4	-18.4
<b>Couderc [23]</b>	64	71.9	EULAR	RTX	ACPA +	68.5	75.9	48.9	7.4	-19.6
<b>Daien [25]</b>	63	69.8	EULAR	RTX	ACPA +	68.5	70.1	62.8	1.6	-5.7
<b>Fabris [27]</b>	30	80.0	EULAR	RTX	ACPA +	68.5	69.5	62.0	1.0	-6.5
<b>Bobbio-Pallavicini [16]</b>	126	65.9	EULAR	TNFi	ACPA -	71.2	78.7	68.7	7.4	-2.5
<b>Chen [21]</b>	48	75.0	EULAR	TNFi	ACPA -	71.2	74.3	70.5	3.1	-0.7
<b>Matsudaira [45]</b>	188	79.3	EULAR	TNFi	ACPA -	71.2	83.8	70.0	12.6	-1.3
<b>Bos [17]</b>	188	77.1	EULAR	TNFi	ACPA +	71.2	71.5	70.5	0.3	-0.8
<b>Braun-Moscovici [18]</b>	30	80.0	$\Delta$ DAS28 1.2	TNFi	ACPA -	71.2	92.5	33.1	21.3	-38.1
<b>Lequerre</b>	33	48.5	$\Delta$ DAS28 1.2	TNFi	ACPA -	71.2	76.7	69.0	5.4	-2.2
<b>Klaasen [40]</b>	89	67.4	$\Delta$ DAS28 1.2	TNFi	ACPA +	71.2	73.6	62.2	2.4	-9.0
<b>Wijbrandts [63]</b>	103	68.0	$\Delta$ DAS28 1.2	TNFi	ACPA +	71.2	83.7	43.1	12.4	-28.1
<b>Pers [50]</b>	165	86.1	EULAR	TCZ	ACPA -	86.1	92.2	82.2	6.1	-3.9
<b>Yukawa [64]</b>	111	73.0	EULAR	TNFi	ANA -	71.2	75.1	59.8	3.9	-11.4
<b>Matsudaira [45]</b>	188	79.3	EULAR	TNFi	ANA +	71.2	72.7	53.8	1.5	-17.4
<b>Couderc [23]</b>	64	71.9	EULAR	RTX	EBV -	68.5	72.8	29.8	4.3	-38.7
<b>Magnusson [43]</b>	35	51.4	$\Delta$ DAS28 1.3	RTX	EBV +	51.4	80.0	30.0	28.6	-21.4
<b>Gottenberg [30]</b>	558	59.1	EULAR	ABA	High IgG	59.1	61.4	57.7	2.2	-1.4
<b>Sellam [59]</b>	208	71.6	EULAR	RTX	High IgG	68.5	78.9	62.7	10.5	-5.8

<b>Kastbom [37]</b>	177	73.4	EULAR	RTX	FCyR-IIIa 158V/F, VF genotype	68.5	79.7	60.0	11.2	-8.5
<b>Quartuccio [52]</b>	212	68.9	EULAR	RTX	FCyR-IIIa 158V/F, VV genotype	68.5	89.0	64.2	20.5	-4.3
<b>Ruysen-Witrand [56]</b>	111	81.1	EULAR	RTX	FCyR-IIIa 158V/F, VV genotype	68.5	86.8	66.1	18.3	-2.4
<b>Canete [20]</b>	73	64.4	EULAR	TNFi	FCyR-IIIa 158V/F, FF genotype	71.2	75.9	68.0	4.7	-3.1
<b>Morales-Lara [46]</b>	30	76.7	EULAR	TNFi	FCyR-IIIa 158V/F, FF genotype	71.2	75.0	69.3	3.9	-1.9
<b>Fabris [28]</b>	158	78.5	EULAR	RTX	IL-6 promotor 174G/C, GC and CC genotype	68.5	73.5	64.8	5.1	-3.7
<b>Davila-Fajardo [26]</b>	191	84.8	EULAR	TNFi	IL-6 promotor 174G/C, GG genotype	71.2	73.7	68.9	2.4	-2.3
<b>Jancic [35]</b>	77	83.1	$\Delta$ DAS28 1.2	TNFi	IL-6 promotor 174G/C, GG genotype	71.2	77.0	68.9	5.8	-2.4
<b>Daïen [25]</b>	63	69.8	EULAR	RTX	TNFA promotor 308G/A, GA genotype	68.5	82.1	63.2	13.6	-5.2
<b>Marotte [44]</b>	198	66.2	ACR 20	TNFi	TNFA promotor 308G/A, GG genotype	58.6	59.9	54.9	1.3	-3.7
<b>Mugnier [48]</b>	53	71.7	$\Delta$ DAS28 1.2	TNFi	TNFA promotor 308G/A, GG genotype	71.2	80.1	41.1	8.9	-30.1
<b>Guis [31]</b>	86	76.7	$\Delta$ DAS28 1.2	TNFi	TNFA promotor 308G/A, GG genotype	71.2	77.8	48.4	6.6	-22.8

ACR50: 20/50% improvement in ACR criteria; APV=added predictive value (i.e. APV+/APV- are the resp. increase and decrease in absolute chance of response when the test-outcome (high or low value of the (bio)marker) is known;  $\Delta$ DAS28 1.2=absolute change in DAS28-score of 1.2; MSM: markers/variables studied multiple times; post probability +/-: the absolute chance of response after test is positive or negative, respectively, when prior probability pooled would be the true setting; TNFi: tumour necrosis factor inhibitor.

Smoking was investigated in two large studies on TNFi [15, 33] and in one on TCZ [50]. In TNFi, smoking was associated with non-response and for TCZ with response. Smoking decreased the chance of EULAR response in TNFi patients from 71.2% (pooled prior probability) with -1.6 to -14.4% (APV-), whereas in TCZ patients, smoking increased the chance of response with 2.8% (APV+).

Presence of erosive disease was reported in six studies, in two using RTX [25, 59] and four using TNFi [17, 18, 40, 63] and predicted response in all studies except for one, based on APV+ and APV- both were 0.0 (as an indirect result of likelihood ratios of 1.0 for both test positive and negative) [18]. Presence of erosive disease increases the chance of EULAR response in RTX patients with 1.9 - 3.4% (APV+), whereas absence of erosive disease decreases the prior probability with 13.9-15.6%. Absence of erosive disease is therefore a relatively important feature in prediction of response to RTX.

Concurrent MTX use shows conflicting results for TNFi and low estimation for other biologicals. Presence of rheumatoid factor (RF) consistently increased the chance of response in RTX patients (APV+ 1.9 to 8.9%) and negativity decreased this chance (APV- -4.1 to -23.7%) based on five studies [23, 25, 27, 53, 58]. Nine studies mentioned RF status in TNFi users [15, 16, 19, 21, 33, 42, 47, 61, 63]. Of these, seven showed an association of RF positivity with increased chance of response (APV+ 0.4 - 12.8), however in the other two studies RF positivity was associated with non-response (APV- -0.1 to -2.0%).

In RTX patients ACPA positivity was consistently associated with response (APV+ 1.1 to 7.5%) and ACPA negativity decreased the chance of response (APV- -5.7 to -19.7%) [23, 25, 27, 53]. ACPA showed inconsistent results for TNFi and only one study for both ABA and TCZ could be included.

Three studies addressed the Fcγ-receptor IIIA 158 V/F polymorphism and response to RTX[37, 52, 56]. The VV-genotype was strongly associated with EULAR response in two of the studies [52, 56] with APV+ of 18.6 and 20.9%, however the third study showed a favour for the VF genotype [37]. In addition, the VV genotype in this third study was associated with the lowest chance of response (44%, data not shown). Two studies addressed polymorphisms of the Fcγ-receptor IIIA and response to TNFi, and showed a weak relationship between the FF genotype and response with APV+ of 2.9 and 4.7% [20, 46].

For TNFi the GG genotype of the IL-6 promotor -174 G>C polymorphism was associated with response (APV+ 2.4 and 5.8%)[26, 35]. The GG genotype in the TNFa 308 G>C polymorphism is associated with response to TNFi therapy (APV+ from 1.3 to 8.9%) and with even higher APV- (thus CC and GC genotype) of -3.7% for ACR20 response and 22.8 and 30.1% for a  $\Delta$ DAS28 >1.2 [31, 44, 48].

### **Markers/variables studied once (MSO)**

All the 65 predictors which were not replicated in additional studies (MSO) are shown in Table 2. This included 9 predictors studied for (non-) response to RTX, 55 predictors for patients treated with a TNFi and one for TCZ. The majority of these predictors consist of gene polymorphisms, proteins, and clinical variables. In general, an APV+ or APV-  $\geq$ 15% is observed more often in the MSO (table 2) as compared to the MSMs (table 1). Twenty-two of the 65 predictors altered the probability of response  $\geq$ 15% in either APV+ or APV- when compared to the prior probability of response. For TNFi the (bio)markers with a APV+  $\geq$ 15% included high IgA, IgM or IgG RF >100 u/ml [39], all RF isotypes positive (>20 u/ml) [39], all CCP2 isotopes positive [39], CC genotype TRAILR1 G/C polymorphism [47], anti-LA positivity [45], the TT genotype of the TNFR1I-codon 196 T/G polymorphism[49], and the RR genotype of Fc $\gamma$ R-2A -131H/R polymorphism[20]. For RTX patients the G/C genotype of TGF $\beta$ 1-25 G/C polymorphism [25] and MRP8/14 serum levels [22] have a APV+  $\geq$ 15%. Predictors for TNFi with a APV-  $\geq$ 15% were BMI <30 [40], anti-Ro negativity [45], IgM-CCP2 positivity [39], Interleukin 17  $\geq$ 40 pg/ml [21], IL-1b >4.84 pg/ml [38], MM genotype of the TNFRSF1B-196 M/R polymorphism [55], GG genotype of the TNFa promotor -238 G/A polymorphism[44], 0-1 copies of the HLA-DRB1 (shared epitope) [36] or (again) TT genotype of the TNFR1I-codon 196 T/G polymorphism [49]. For RTX, predictors with a APV-  $\geq$ 15% were being seropositive for RF or ACPA[59], Immunoglobulins, <6 or >12gm/litre [23], TC genotype of TGF- $\beta$ 1-10 T/C polymorphism [25] and (again) MRP8/14 serum levels [22].

**Table 2:** Overview of predictors studied once.

Author (et al.)	Class of biological	Predictors	Total no.	Prior prob per study, %	PPV, %	NPV, %	APV+, %	APV-, %
Choi [22]	RTX	MRP8/14 serum levels, high	24	54.2	91.7	83.3	37.5	-37.5
Couderc [23]	RTX	Immunoglobulins, <6 or >12gm/litre	62	74.2	80.4	43.8	6.2	-17.9
Daïen [25]	RTX	TGFβ1-10 T/C, TC genotype	63	49.2	61.4	78.9	12.2	-28.2
Daïen [25]	RTX	TGFβ1-25 G/C, GC genotype	63	69.8	100.0	36.5	30.2	-6.4
Fabris [27]	RTX	Anti-MCV+	30	80.0	80.8	25.0	0.8	-5.0
Ruyssen-Witrand [57]	RTX	BAFF-871C/T, CC genotype	115	80.9	92.3	25.0	11.4	-5.9
Sellam [59]	RTX	RF or ACPA +	208	71.6	76.5	50.0	4.8	-21.6
Sellam [59]	RTX	DAS28-CRP >5.1	208	71.6	76.3	42.3	4.6	-13.9
Sellam [59]	RTX	>2.69 gm/liter IgA	208	71.6	78.6	33.1	6.9	-4.7
Buch [19]*	TNFi	CRP, ≥20% reduction by week 2	207	51.7	57.0	86.0		
Canete [20]	TNFi	FcyR-2A -131H/R, RR-genotype	81	39.5	60.0	67.2	20.5	-6.7
Choi [22]	TNFi	MRP8/14 serum levels, high	146	75.3	89.0	38.4	13.7	-13.7
Coulthard [24]	TNFi	MAKK14 (rs916344) C/G, genotype 1+2 vs 0	901	54.9	57.0	53.4	2.0	-8.3
Coulthard [24]	TNFi	MAP2K6 (rs11656130) T/G, genotype 1+2 vs 0	918	29.6	31.6	78.5	2.0	-8.1
Coulthard [24]	TNFi	RPS6KA4 (rs475032) G/C, genotype 1+2 vs 0	803	43.8	45.2	60.9	1.4	-4.7
Coulthard [24]	TNFi	RPS6KA5 (rs1286112) C/G, genotype 0 vs 1+2	920	52.2	52.8	50.3	0.6	-2.5
Coulthard [24]	TNFi	MAP2K6 (rs2716191) T/C, genotype 1+2 vs 0	805	34.2	34.1	65.6	0.3	-0.1
Coulthard [24]	TNFi	RPS6KA5 (rs1286076) C/T, genotype 0 vs 1+2	900	52.9	52.8	46.6	0.5	-0.1
Coulthard [24]	TNFi	MAPKAPK2 (rs4240847) C/A, genotype 1+2 vs 0	918	47.9	47.7	51.1	1.0	-0.2
Fisher [29]	TNFi	Anti-CEP antibodies+	450	78.7	79.3	22.4	0.7	-1.1
Fisher [29]	TNFi	Anti-FBP antibodies+	436	78.9	81.5	22.0	2.6	-0.9
Fisher [29]	TNFi	Anti-cVIM antibodies+	436	78.9	79.3	21.4	0.4	-0.3
Hassan [32]	TNFi	sIL-6R A/C, AA+AC genotype	999	80.1	80.1	20.1	0.0	-0.2
Jamnitski [34]	TNFi	Anti-IFX antibodies before start ETN	89	79.8	91.5	33.3	11.7	-13.1
Kang [36]	TNFi	HLA-DRB1 (SE), 0 or 1 copy	66	86.4	89.5	33.3	3.1	-19.7
Kang [36]	TNFi	TNFA promotor 857C/T, CT+TT genotype	70	85.7	96.3	20.9	10.6	-6.6
Kayakabe [38]*	TNFi	IL-1b, >4.84 pg/ml	41	78.0	92.6	50.0	14.6	28.0
Klaasen [40]	TNFi	BMI < 30	79	75.9	82.8	53.3	6.9	-29.3
Klaasen [39]	TNFi	IgM-CCP2+	101	68.3	77.8	47.4	9.5	-15.7
Klaasen [39]	TNFi	IgG-CCP2+	101	68.3	72.7	45.8	4.4	-14.2
Klaasen [39]	TNFi	All CCP2-isotopes +	101	68.3	88.1	45.8	19.8	-14.1
Klaasen [41]	TNFi	lymphocyte aggregates present in syn. tissue	97	75.3	83.6	35.7	8.4	-11.0
Klaasen [39]	TNFi	IgM RF, >100 u/ml	101	68.3	85.7	38.4	17.4	-6.7
Klaasen [39]	TNFi	IgA-CCP2+	101	68.3	73.1	36.7	4.8	-5.1

<b>Klaasen [39]</b>	TNFi	All RF-isotopes +	101	68.3	84.2	35.4	15.9	-3.7
<b>Klaasen [39]</b>	TNFi	IgA RF, >100 u/ml	101	68.3	100.0	33.7	31.7	-2.0
<b>Klaasen [39]</b>	TNFi	IgG RF, >100 u/ml	101	68.3	100.0	33.0	31.7	-1.3
<b>Marotte [44]</b>	TNFi	TNFa promotor 238 G/A, GG genotype	198	66.2	67.2	55.6	1.0	-21.7
<b>Marotte [44]</b>	TNFi	IL1b -3954 C/T, C/C genotype	198	66.2	71.3	41.0	5.1	-7.1
<b>Marotte [44]</b>	TNFi	IL1-RN -2018 T/C, T/T genotype	198	66.2	68.5	36.7	2.4	-2.8
<b>Marotte [44]</b>	TNFi	Shared epitope, carrier	198	66.2	66.7	35.1	0.5	-1.2
<b>Matsudaira [45]</b>	TNFi	Anti-Ro negativity	188	79.3	83.4	41.9	4.2	-21.2
<b>Matsudaira [45]</b>	TNFi	Anti-LA positivity	188	79.3	100.0	21.1	20.7	-0.3
<b>Morales-Lara [47]</b>	TNFi	TRAILR1 G/C, CC genotype	89	62.9	82.6	43.9	19.7	-6.9
<b>Morales-Lara [47]</b>	TNFi	TNFR1A A/G, AA genotype	89	62.9	66.7	38.7	3.7	-1.6
<b>Pers [49]</b>	TNFi	LTA+ 720 C/A, CC genotype	15	66.7	80.0	40.0	13.3	-6.7
<b>Pers [49]</b>	TNFi	IL-10-1087 A/G, AA genotype	15	66.7	75.0	36.4	8.3	-3.0
<b>Pers [49]</b>	TNFi	TNFRII-codon 196 T/G, TT genotype	15	66.7	90.0	80.0	23.3	-46.7
<b>Potter [51]</b>	TNFi	PTGS2 G/A, GG genotype	903	80.2	81.2	30.1	1.0	-10.3
<b>Potter [51]</b>	TNFi	NFkBIB rs3136645 T/C, TT genotype	902	80.2	82.3	24.6	2.2	-4.8
<b>Potter [51]</b>	TNFi	TLR-2 C/G, CC genotype	901	79.9	81.8	24.1	1.9	-4.0
<b>Potter [51]</b>	TNFi	NFkBIB rs9403, G/C, GG genotype	893	80.4	83.9	22.3	3.5	-2.7
<b>Potter [51]</b>	TNFi	IRAK-3 T/A, TA genotype	901	80.1	84.4	21.6	4.2	-1.7
<b>Potter [51]</b>	TNFi	CHUK GG genotype	883	80.0	83.3	21.5	3.4	-1.5
<b>Potter [51]</b>	TNFi	MyD88 A/G, AG genotype	902	80.2	83.8	21.1	3.7	-1.3
<b>Potter [51]</b>	TNFi	TLR-10/1/6 A/C, CC genotype	894	79.9	86.5	20.9	6.6	-0.8
<b>Potter [51]</b>	TNFi	IKKBK A/C, , AC genotype	901	80.5	91.4	20.3	10.9	-0.8
<b>Radstake [54]</b>	TNFi	MIF 173G/C, C-allele carrier	90	71.1	81.3	34.5	10.1	-5.6
<b>Radstake [54]</b>	TNFi	MIF (CATT)n repeat, CATT-7 carrier	90	71.1	81.0	31.9	9.8	-3.0
<b>Rooryck [55]</b>	TNFi	TNFRSF1B-196 M/R, MM genotype	78	56.4	68.0	64.3	11.6	-20.7
<b>Rooryck [55]</b>	TNFi	FCGR3A-212V/F, FF genotype	78	56.4	56.7	43.8	0.3	-0.2
<b>Straub [60]*†</b>	TNFi	Ratio serum cortisol/serum ACTH	38	50.0	90.0			
<b>Toonen [62]</b>	TNFi	TNFSF1b 676T/G, GT genotype	234	72.2	76.0	30.4	3.8	-2.7
<b>Pers [50]</b>	TCZ	CRP>10mg/l	165	86.1	93.1	19.4	7.0	-5.4

\* = no actual numbers of test positive/negative versus responders/non-responders extractable from the main article. Although information on PPV or NPV was available.

† = data extracted from validation cohort

APV=added predictive value (i.e. APV+/APV- are the resp. increase and decrease in absolute chance of response when the test-outcome (high or low value of the (bio)marker) is known, NPV=negative predictive value (i.e. the proportion of test negatives which are non-responders (true negatives) among all test negatives), PPV=positive predictive value (i.e. the proportion of test positives which are responders (true positives) among all test positives).

## **Multibiomarker models**

All extracted multibiomarker models, except those with adjustments for clinical parameters are presented as a supplementary table (see Supplementary Table 2).

## **DISCUSSION**

In our extensive literature search and subsequent analyses we were able to compile and compare many (bio)markers. Since none of the MSMs met the predefined criterion of a consistent increase or decrease in added predicted value of  $\geq 15\%$ , this review was not able to find clinically useful baseline (bio)markers that can be used in individually tailored treatment. However, among the MSMs with lower predictive values, there is a possible role for RF and ACPA positivity in RTX and the GG genotype of the TNF $\alpha$  promotor -308 G>C polymorphism for TNFi treatment, because these markers showed a consistent direction in each study. Several of the markers studied once (MSO) yielded high added predictive values and are therefore of clinical interest. However, these high values could be a result of publication bias and might therefore follow a 'regression to the mean' (if not contradictory results) when being replicated in other studies, analogous to the MSMs. To implement any of the studied MSO in daily practice, extensive validation is a prerequisite. The risk of bias assessment revealed that most of the included studies have a moderate risk of bias. However, because none of the predictors showed a consistent increase or decrease in added predicted value of  $\geq 15\%$ , this finding did not influence our interpretation of the study. The allocation of a moderate or high risk was usually a result from the exclusion of certain patients from all eligible patients or inadequate dealing with missing data, both of which increase bias and negatively influences the clinical representability of the results. A focus for future researchers to these important aspects is highly recommended, since this would increase the validity and reproducibility of prediction studies. Among the included studies, RF and ACPA positivity predicted modestly, however consistently, response to RTX with an APV+ of 1.9-8.9% and 1.1-7.5% respectively. These findings are in line with a meta-analysis on 4 placebo controlled phase II or III clinical trials[72], in which it was concluded that seropositive patients (RF and/or ACPA) respond better than sero-negative patients treated with RTX. The TNF- $\alpha$  promotor -

308 G/A polymorphism was investigated in four of the included studies. In these studies, GG genotype predicts response to TNFi. In a meta-analysis from Zeng et al. [73], it is indicated that the TNF- $\alpha$  308 G allele is a predictor for response to TNFi treatment. However, in 2 other meta-analyses [74, 75] contradictory results are presented indicating no correlation with response. The added value in clinical practice of this predictor remains therefore questionable.

The results of this systematic review should be interpreted taking the drawbacks into account. The first and most important issue is that we were restricted to a selection of all relevant studies for data extraction. A decision for individualized treatment cannot be based on p-values alone, since they merely inform on the probability that a hypothesis is correct and not on the magnitude of the effect. Therefore, our focus was on calculating interpretable outcomes like (difference in) pre- and post-test probabilities. In the majority of the excluded studies, these values could not be extracted or calculated. The corresponding authors of these excluded articles were not contacted to retrieve data in the correct format, which is a potential drawback of this study. We also excluded all multivariable models containing clinical variables, since comparisons between models and extraction of quantifiable data from these models is not possible. Because many studies had to be excluded due to the inability to extract data (n=72), a certain amount of bias will be present. Regardless of the strict selection criteria, 57 studies provided clinically applicable data which yielded many predictors. Multiple testing and/or absence of internal/external validation, will add to the increased the chance on false positive results in the included studies. Alongside with reporting- and publication bias which both play a significant role in the integrity of medical research, these issues could contribute to less reliable results. Moreover, when all proposed predictors would be replicated in other cohorts and all results (also the negative) would be published, it is likely that estimations of predictive values decrease even further. We adjusted partially for this by searching in each included study for variables with quantifiable information that could be included as MSM. This additional data resulted in a more valid estimation of the true value of these markers in clinical practice.

Nowadays, immunogenicity has emerged as an important theme on prediction of response to drugs [76, 77]. For example, anti-drug antibodies and serum drug level

measurements have been explored and presented in the literature as explanatory markers for (non-)response [78-81]. However, most of these studies did not fulfil the criteria in our study since they presented results on continuous values or anti-drug antibodies were measured later than four weeks after initiation of therapy. The added value in clinical practice is probably limited when it can only be measured after that time, because early clinical response to treatment can already be assessed. These (bio)markers are therefore more interesting in etiologic sense and provide insight in possible other predictors.

## **CONCLUSION**

Many studies have investigated predictors of response to biologicals in rheumatoid arthritis. However, this review was not able to find clinically useful baseline (bio)markers that can be used in individually tailored treatment. The challenge now is to design studies to validate all explored and promising findings to make (combinations of) these (bio)markers relevant to clinical practice. In this future research, it is preferable to have validation cohorts reflecting clinical practice, a test with a clear interpretable outcome and/or cut-off value and results related and based on absolute chances of (non-)response.

## REFERENCES

1. Gabriel SE, Michaud K. Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. *Arthritis Res Ther* 2009;11 3: 229.
2. Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73 3: 492-509.
3. Keystone EC, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum* 2004;50 5: 1400-11.
4. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med* 1999;340 4: 253-9.
5. Furst DE. The risk of infections with biologic therapies for rheumatoid arthritis. *Semin Arthritis Rheum* 2010;39 5: 327-46.
6. Kievit W, Adang EM, Franssen J, Kuper HH, van de Laar MA, Jansen TL et al. The effectiveness and medication costs of three anti-tumour necrosis factor alpha agents in the treatment of rheumatoid arthritis from prospective clinical practice data. *Ann Rheum Dis* 2008;67 9: 1229-34.
7. Mass RD, Press MF, Anderson S, Cobleigh MA, Vogel CL, Dybdal N et al. Evaluation of clinical outcomes according to HER2 detection by fluorescence in situ hybridization in women with metastatic breast cancer treated with trastuzumab. *Clin Breast Cancer* 2005;6 3: 240-6.
8. Ross JS, Slodkowska EA, Symmans WF, Pusztai L, Ravdin PM, Hortobagyi GN. The HER-2 receptor and breast cancer: ten years of targeted anti-HER-2 therapy and personalized medicine. *Oncologist* 2009;14 4: 320-68.
9. van Gestel AM, Prevoost ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996;39 1: 34-40.
10. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;38 6: 727-35.
11. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;158 4: 280-6.
12. Mistry K, Cable G. Meta-analysis of prostate-specific antigen and digital rectal examination as screening tests for prostate carcinoma. *J Am Board Fam Pract* 2003;16 2: 95-101.
13. Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J Clin Oncol* 2002;20 3: 719-26.
14. Seidman AD, Fornier MN, Esteva FJ, Tan L, Kaptain S, Bach A et al. Weekly trastuzumab and paclitaxel therapy for metastatic breast cancer with analysis of efficacy by HER2 immunophenotype and gene amplification. *J Clin Oncol* 2001;19 10: 2587-95.
15. Abhishek A, Butt S, Gadsby K, Zhang W, Deighton CM. Anti-TNF-(alpha) agents are less effective for the treatment of rheumatoid arthritis in current smokers. *Journal of Clinical Rheumatology* 2010;16 1: 15-8.
16. Bobbio-Pallavicini F, Caporali R, Alpini C, Avalle S, Epis OM, Klersy C et al. High IgA rheumatoid factor levels are associated with poor clinical response to tumour necrosis factor (alpha) inhibitors in rheumatoid arthritis. *Annals of the Rheumatic Diseases* 2007;66 3: 302-7.
17. Bos WH, Bartelds GM, Wolbink GJ, de Koning MH, van de Stadt RJ, van Schaardenburg D et al. Differential response of the rheumatoid factor and anticitrullinated protein antibodies during adalimumab treatment in patients with rheumatoid arthritis. *J Rheumatol*. 2008;1972-7.
18. Braun-Moscovici Y, Markovits D, Zinder O, Schapira D, Rozin A, Ehrenburg M et al. Anticyclic citrullinated protein antibodies as a predictor of response to anti-tumor necrosis factor-alpha therapy in patients with rheumatoid arthritis. *J Rheumatol* 2006;33 3: 497-500.

19. Buch MH, Seto Y, Bingham SJ, Bejarano V, Bryer D, White J et al. C-reactive protein as a predictor of infliximab treatment outcome in patients with rheumatoid arthritis: defining subtypes of nonresponse and subsequent response to etanercept. *Arthritis Rheum* 2005;52 1: 42-8.
20. Canete JD, Suarez B, Hernandez MV, Sanmarti R, Rego I, Celis R et al. Influence of variants of Fc gamma receptors IIA and IIIA on the American College of Rheumatology and European League Against Rheumatism responses to anti-tumour necrosis factor alpha therapy in rheumatoid arthritis. *Ann Rheum Dis*. 2009;1547-52.
21. Chen DY, Chen YM, Chen HH, Hsieh CW, Lin CC, Lan JL. Increasing levels of circulating Th17 cells and interleukin-17 in rheumatoid arthritis patients with an inadequate response to anti-TNF-alpha therapy. *Arthritis Res Ther*. 2011;R126.
22. Choi YK, Herenius MJ, Wijbrandts CA, Roth J, Foll D, Gerlag DM et al. MRP8/14 serum complexes as predictor of response to biological treatments in rheumatoid arthritis. *Annals of the Rheumatic Diseases* 2013;72: 499-505.
23. Couderc M, Mathieu S, Pereira B, Glace B, Soubrier M. Predictive factors of rituximab response in rheumatoid arthritis: results from a French university hospital. *Arthritis Care Res (Hoboken)* 2013;65 4: 648-52.
24. Coulthard LR, Taylor JC, Eyre S, Robinson JI, Wilson AG, Isaacs JD et al. Genetic variants within the MAP kinase signaling network and anti-TNF treatment response in Rheumatoid arthritis patients. *Annals of the Rheumatic Diseases* 2011;70 1: 98-103.
25. Daien CI, Fabre S, Rittore C, Soler S, Daien V, Tejedor G et al. TGF beta1 polymorphisms are candidate predictors of the clinical response to rituximab in rheumatoid arthritis. *Joint Bone Spine*. 2012;471-5.
26. Davila-Fajardo CL, Marquez A, Pascual-Salcedo D, Moreno Ramos MJ, Garcia-Portales R, Magro C et al. Confirmation of -174G/C interleukin-6 gene promoter polymorphism as a genetic marker predicting antitumor necrosis factor treatment outcome. *Pharmacogenet Genomics* 2014;24 1: 1-5.
27. Fabris M, de Vita S, Blasone N, Visentini D, Pezzarini E, Pontarini E et al. Serum levels of anti-CCP antibodies, anti-MCV antibodies and RF iga in the follow-up of patients with rheumatoid arthritis treated with rituximab. *Autoimmunity Highlights* 2010;1 2: 87-94.
28. Fabris M, Quartuccio L, Lombardi S, Saracco M, Atzeni F, Carletto A et al. The CC homozygosis of the -174G>C IL-6 polymorphism predicts a lower efficacy of rituximab therapy in rheumatoid arthritis. *Autoimmun Rev*. 2012;315-20.
29. Fisher BA, Plant D, Lundberg K, Charles P, Barton A, Venables PJ. Heterogeneity of anticitrullinated peptide antibodies and response to anti-tumor necrosis factor agents in rheumatoid arthritis. *J Rheumatol* 2012;39 5: 929-32.
30. Gottenberg JE, Ravaud P, Cantagrel A, Combe B, Flipo RM, Schaevebeke T et al. Positivity for anti-cyclic citrullinated peptide is associated with a better response to abatacept: data from the 'Orencia and Rheumatoid Arthritis' registry. *Ann Rheum Dis*. 2012;1815-9.
31. Guis S, Balandraud N, Bouvenot J, Auger I, Toussiroit E, Wendling D et al. Influence of -308 A/G polymorphism in the tumor necrosis factor alpha gene on etanercept treatment in rheumatoid arthritis. *Arthritis Rheum*. 2007;1426-30.
32. Hassan B, Maxwell JR, Hyrich KL, Barton A, Worthington J, Isaacs JD et al. Genotype at the sIL-6R A358C polymorphism does not influence response to anti-TNF therapy in patients with rheumatoid arthritis. *Rheumatology (Oxford)*. 2010;43-7.
33. Hyrich KL, Watson KD, Silman AJ, Symmons DP. 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: 1558-65.
34. Jamnitski A, Kriekaert CL, Nurmohamed MT, Hart MH, Dijkmans BA, Aarden L et al. Patients non-responding to etanercept obtain lower etanercept concentrations compared with responding patients. *Ann Rheum Dis*. 2011;88-91.
35. Jancic I, Arsenovic-Ranin N, Sefik-Bukilica M, Zivojinovic S, Damjanov N, Spasovski V et al. -174G/C interleukin-6 gene promoter polymorphism predicts therapeutic response to etanercept in rheumatoid arthritis. *Rheumatol Int*. 2013;1481-6.
36. Kang CP, Lee KW, Yoo DH, Kang C, Bae SC. The influence of a polymorphism at position -857 of the tumour necrosis factor (alpha) gene on clinical response to etanercept therapy in rheumatoid arthritis. *Rheumatology* 2005;44 4: 547-52.
37. Kastbom A, Coster L, Arlestig L, Chatzidionysiou A, Van Vollenhoven RF, Padyukov L et al. Influence of FCGR3A genotype on the therapeutic response to rituximab in rheumatoid arthritis: an

- observational cohort study. *BMJ Open* 2012;2 5: e001524.
38. Kayakabe K, Kuroiwa T, Sakurai N, Ikeuchi H, Kadiombo AT, Sakairi T et al. Interleukin-1beta measurement in stimulated whole blood cultures is useful to predict response to anti-TNF therapies in rheumatoid arthritis. *Rheumatology (Oxford)* 2012;51 9: 1639-43.
  39. Klaasen R, Cantaert T, Wijbrandts CA, Teitsma C, Gerlag DM, Out TA et al. The value of rheumatoid factor and anti-citrullinated protein antibodies as predictors of response to infliximab in rheumatoid arthritis: an exploratory study. *Rheumatology (Oxford)* 2011;50 8: 1487-93.
  40. Klaasen R, Wijbrandts CA, Gerlag DM, Tak PP. Body mass index and clinical response to infliximab in rheumatoid arthritis. *Arthritis Rheum* 2011;63 2: 359-64.
  41. Klaasen R, Thurlings RM, Wijbrandts CA, van Kuijk AW, Baeten D, Gerlag DM et al. The relationship between synovial lymphocyte aggregates and the clinical response to infliximab in rheumatoid arthritis: a prospective study. *Arthritis Rheum* 2009;60 11: 3217-24.
  42. Lequerre T, Gauthier-Jauneau AC, Bansard C, Derambure C, Hiron M, Vittecoq O et al. Gene profiling in white blood cells predicts infliximab responsiveness in rheumatoid arthritis. *Arthritis Res Ther* 2006;8 4: R105.
  43. Magnusson M, Brisslert M, Zendjanchi K, Lindh M, Bokarewa MI. Epstein-Barr virus in bone marrow of rheumatoid arthritis patients predicts response to rituximab treatment. *Rheumatology (Oxford)*. 2010:1911-9.
  44. Marotte H, Pallot-Prades B, Grange L, Tebib J, Gaudin P, Alexandre C et al. The shared epitope is a marker of severity associated with selection for, but not with response to, infliximab in a large rheumatoid arthritis population. *Ann Rheum Dis*. 2006:342-7.
  45. Matsudaira R, Tamura N, Sekiya F, Ogasawara M, Yamanaka K, Takasaki Y. Anti-Ro/SSA antibodies are an independent factor associated with an insufficient response to tumor necrosis factor inhibitors in patients with rheumatoid arthritis. *J Rheumatol*. 2011:2346-54.
  46. Morales-Lara MJ, Conesa-Zamora P, Garcia-Simon MS, Pedrero F, Santaclara V, Perez-Guillermo M et al. Association between the FCGR3A V158F polymorphism and the clinical response to infliximab in rheumatoid arthritis and spondyloarthritis patients. *Scand J Rheumatol* 2010;39 6: 518-20.
  47. Morales-Lara MJ, Canete JD, Torres-Moreno D, Hernandez MV, Pedrero F, Celis R et al. Effects of polymorphisms in TRAILR1 and TNFR1A on the response to anti-TNF therapies in patients with rheumatoid and psoriatic arthritis. *Joint Bone Spine*. 2012:591-6.
  48. Mugnier B, Balandraud N, Darque A, Roudier C, Roudier J, Revirion D. Polymorphism at position -308 of the tumor necrosis factor alpha gene influences outcome of infliximab therapy in rheumatoid arthritis. *Arthritis Rheum*. 2003:1849-52.
  49. Pers YM, Cadart D, Rittore C, Ravel P, Daien V, Fabre S et al. TNFR2 polymorphism is associated with response to TNF blockers in rheumatoid arthritis patients seronegative for ACPA. *Joint Bone Spine* 2014;81 4: 370-2.
  50. Pers YM, Fortunet C, Constant E, Lambert J, Godfrin-Valnet M, De JA et al. Predictors of response and remission in a large cohort of rheumatoid arthritis patients treated with tocilizumab in clinical practice. *Rheumatology (Oxford)* 2014;53 1: 76-84.
  51. Potter C, Gibbons LJ, Bowes JD, Cordell HJ, Hyrich K, Isaacs JD et al. Polymorphisms spanning the TNFR2 and TACE genes do not contribute towards variable anti-TNF treatment response. *Pharmacogenet Genomics*. 2010:338-41.
  52. Quartuccio L, Fabris M, Pontarini E, Salvin S, Zabotti A, Benucci M et al. The 158VV Fcgamma receptor 3A genotype is associated with response to rituximab in rheumatoid arthritis: results of an Italian multicentre study. *Ann Rheum Dis*. 2013:716-21.
  53. Quartuccio L, Fabris M, Salvin S, Atzeni F, Saracco M, Benucci M et al. Rheumatoid factor positivity rather than anti-CCP positivity, a lower disability and a lower number of anti-TNF agents failed are associated with response to rituximab in rheumatoid arthritis. *Rheumatology* 2009;48 12: 1557-9.
  54. Radstake TR, Fransen J, Toonen EJ, Coenen MJ, Eijsbouts AE, Donn R et al. Macrophage migration inhibitory factor polymorphisms do not predict therapeutic response to glucocorticoids or to tumour necrosis factor alpha-neutralising treatments in rheumatoid arthritis. *Ann Rheum Dis*. 2007:1525-30.
  55. Rooryck C, Barnette T, Richez C, Laleye A, Arveiler B, Schaeverbeke T. Influence of FCGR3A-V212F and TNFRSF1B-M196R genotypes in patients with rheumatoid arthritis treated with infliximab therapy. *Clin Exp Rheumatol*. 2008:340-2.
  56. Ruysen-Witrand A, Rouanet S, Combe B, Dougados M, Le Loet X, Sibilia J et al. Fc(gamma) receptor type IIIA polymorphism influences treatment outcomes in patients

- with rheumatoid arthritis treated with rituximab. *Annals of the Rheumatic Diseases* 2012;71 6: 875-7.
57. Ruysen-Witrand A, Rouanet S, Combe B, Dougados M, Le Loet X, Sibilia J et al. Association between -871C>T promoter polymorphism in the B-cell activating factor gene and the response to rituximab in rheumatoid arthritis patients. *Rheumatology (Oxford)*. 2013;636-41.
  58. Soliman MM, Hyrich KL, Lunt M, Watson KD, Symmons DP, Ashcroft DM. Rituximab or a second anti-tumor necrosis factor therapy for rheumatoid arthritis patients who have failed their first anti-tumor necrosis factor therapy? Comparative analysis from the British Society for Rheumatology Biologics Register. *Arthritis Care Res (Hoboken)* 2012;64 8: 1108-15.
  59. Sellam J, Hendel-Chavez H, Rouanet S, Abbed K, Combe B, Le L, X et al. B cell activation biomarkers as predictive factors for the response to rituximab in rheumatoid arthritis: a six-month, national, multicenter, open-label study. *Arthritis Rheum* 2011;63 4: 933-8.
  60. Straub RH, Pongratz G, Cutolo M, Wijnbrandts CA, Baeten D, Fleck M et al. Increased cortisol relative to adrenocorticotropic hormone predicts improvement during anti-tumor necrosis factor therapy in rheumatoid arthritis. *Arthritis Rheum*. 2008;976-84.
  61. Stuhlmüller B, Haupl T, Hernandez MM, Grutzkau A, Kuban RJ, Tandon N et al. CD11c as a transcriptional biomarker to predict response to anti-TNF monotherapy with adalimumab in patients with rheumatoid arthritis. *Clin Pharmacol Ther* 2010;87 3: 311-21.
  62. Toonen EJM, Coenen MJH, Kievit W, Fransen J, Eijsbouts AM, Scheffer H et al. The tumour necrosis factor receptor superfamily member 1b 676T>G polymorphism in relation to response to infliximab and adalimumab treatment and disease severity in rheumatoid arthritis. *Annals of the Rheumatic Diseases* 2008;67 8: 1174-7.
  63. Wijnbrandts CA, Dijkgraaf MGW, Kraan MC, Vinkenoog M, Smeets TJ, Dinant H et al. The clinical response to infliximab in rheumatoid arthritis is in part dependent on pretreatment tumour necrosis factor (alpha) expression in the synovium. *Annals of the Rheumatic Diseases* 2008;67 8: 1139-44.
  64. Yukawa N, Fujii T, Kondo-Ishikawa S, Yoshifuji H, Kawabata D, Nojima T et al. Correlation of antinuclear antibody and anti-double-stranded DNA antibody with clinical response to infliximab in patients with rheumatoid arthritis: a retrospective clinical study. *Arthritis Res Ther*. 2011;R213.
  65. Bansard C, Lequerre T, Derambure C, Vittecoq O, Hiron M, Daragon A et al. Gene profiling predicts rheumatoid arthritis responsiveness to IL-1Ra (anakinra). *Rheumatology (Oxford)* 2011;50 2: 283-92.
  66. Bazzichi L, Rossi P, Giacomelli C, De FF, Bobbio-Pallavicini F, Rossi A et al. A proposal of simple calculation (ERI calculator) to predict the early response to TNF-alpha blockers therapy in rheumatoid arthritis. *Rheumatol Int* 2012;32 2: 349-56.
  67. Bienkowska JR, Dalgin GS, Batliwalla F, Allaire N, Roubenoff R, Gregersen PK et al. Convergent Random Forest predictor: methodology for predicting drug response from genome-scale data applied to anti-TNF response. *Genomics*. 2009;423-232.
  68. Julia A, Erra A, Palacio C, Tomas C, Sans X, Barcelo P et al. An eight-gene blood expression profile predicts the response to infliximab in rheumatoid arthritis. *PLoS One* 2009;4 10: e7556.
  69. Lequerre T, Jouen F, Brazier M, Clayssens S, Klemmer N, Menard JF et al. Autoantibodies, metalloproteinases and bone markers in rheumatoid arthritis patients are unable to predict their responses to infliximab. *Rheumatology (Oxford)* 2007;46 3: 446-53.
  70. Tanino M, Matoba R, Nakamura S, Kameda H, Amano K, Okayama T et al. Prediction of efficacy of anti-TNF biologic agent, infliximab, for rheumatoid arthritis patients using a comprehensive transcriptome analysis of white blood cells. *Biochem Biophys Res Commun* 2009;387 2: 261-5.
  71. Thurlings RM, Boumans M, Tekstra J, van Roon JA, Vos K, van Westing DM et al. Relationship between the type I interferon signature and the response to rituximab in rheumatoid arthritis patients. *Arthritis Rheum* 2010;62 12: 3607-14.
  72. Isaacs JD, Cohen SB, Emery P, Tak PP, Wang J, Lei G et al. Effect of baseline rheumatoid factor and anticitrullinated peptide antibody serotype on rituximab clinical response: a meta-analysis. *Ann Rheum Dis* 2013;72 3: 329-36.
  73. Zeng Z, Duan Z, Zhang T, Wang S, Li G, Gao J et al. Association between tumor necrosis factor-alpha (TNF-alpha) promoter -308 G/A and response to TNF-alpha blockers in rheumatoid arthritis: a meta-analysis. *Mod Rheumatol* 2013;23 3: 489-95.
  74. Lee YH, Ji JD, Bae SC, Song GG. Associations between tumor necrosis factor-alpha (TNF-alpha) -308 and -238 G/A polymorphisms and

- shared epitope status and responsiveness to TNF-alpha blockers in rheumatoid arthritis: a metaanalysis update. *J Rheumatol* 2010;37 4: 740-6.
75. Pavy S, Toonen EJ, Miceli-Richard C, Barrera P, van Riel PL, Criswell LA et al. Tumour necrosis factor alpha -308G->A polymorphism is not associated with response to TNFalpha blockers in Caucasian patients with rheumatoid arthritis: systematic review and meta-analysis. *Ann Rheum Dis* 2010;69 6: 1022-8.
  76. Mandema JW, Salinger DH, Baumgartner SW, Gibbs MA. A dose-response meta-analysis for quantifying relative efficacy of biologics in rheumatoid arthritis. *Clin Pharmacol Ther* 2011;90 6: 828-35.
  77. Maneiro JR, Salgado E, Gomez-Reino JJ. Immunogenicity of monoclonal antibodies against tumor necrosis factor used in chronic immune-mediated Inflammatory conditions: systematic review and meta-analysis. *JAMA Intern Med* 2013;173 15: 1416-28.
  78. Bartelds GM, Wijbrandts CA, Nurmohamed MT, Stapel S, Lems WF, Aarden L et al. Anti-infliximab and anti-adalimumab antibodies in relation to response to adalimumab in infliximab switchers and anti-tumour necrosis factor naive patients: a cohort study. *Ann Rheum Dis*. 2010;817-21.
  79. Krintel SB, Grunert VP, Hetland ML, Johansen JS, Rothfuss M, Palermo G et al. The frequency of anti-infliximab antibodies in patients with rheumatoid arthritis treated in routine care and the associations with adverse drug reactions and treatment failure. *Rheumatology (Oxford)*. 2013:1245-53.
  80. Jamnitski A, Bartelds GM, Nurmohamed MT, van Schouwenburg PA, van Schaardenburg D, Stapel SO et al. The presence or absence of antibodies to infliximab or adalimumab determines the outcome of switching to etanercept. *Ann Rheum Dis*. 2012:284-8.
  81. Wolbink GJ, Voskuyl AE, Lems WF, de GE, Nurmohamed MT, Tak PP et al. Relationship between serum trough infliximab levels, pretreatment C reactive protein levels, and clinical response to infliximab treatment in patients with rheumatoid arthritis. *Ann Rheum Dis* 2005;64 5: 704-7.

**Supplementary Text 1:** Search syntax

(rheumatoid AND arthritis) OR RA

---

(biologics OR biological OR biologicals OR (monoclonal and antibodies) OR (anti AND tnfalpa) OR anti-TNF OR antitnf OR (TNF AND antagonist) OR (TNF AND antagonists) OR (TNF AND inhibitor) OR (TNF AND inhibitors) OR Adalimumab OR Etanercept OR Infliximab OR Golimumab OR certolizumab OR Anakinra OR interleukin OR Tocilizumab OR (anti AND il\*) OR Rituximab OR anticd\* OR (anti AND cd\*) OR (B AND cell AND depleting) OR (B AND cell AND depletion) OR Abatacept)

---

(biomarker OR biomarkers OR marker OR markers OR predicts OR predicted OR predicting OR predictive OR prognostic OR prognosis OR (risk AND factor) OR association OR associations OR associates OR variant OR polymorphism OR polymorphisms OR receptor OR personalized OR differentiate OR differentiates OR differentiated OR differentiating OR (gene AND expression) OR (rheumatoid AND factor) OR ACPA OR antibodies)

---

(response OR responses OR responder OR responded OR responders OR responsiveness OR ACR-20 OR ACR-50 OR ACR-70 OR DAS28)

**Supplementary Table 1:** Risk of bias assessment for included studies (n=56). Items “5. Study Confounding” and “6. Statistical Analysis and Reporting” (not shown) were not scored because we extracted unadjusted and unanalyzed data from the studies (with the exception of the multivariable biomarker models). For item 1. study participation, all studies that included only biological naïve patients were scored ‘moderate’ because it is doubtful if these patients reflect clinical practice.

<b>Author (et al.)</b>	<b>1. Study Participation</b>	<b>2. Study Attrition</b>	<b>3. Progn. Factor Measurement</b>	<b>4. Outcome Measurement</b>
<b>Abhishek [15]</b>	moderate	high	high	low
<b>Bansard [65]</b>	moderate	low	moderate	low
<b>Bazzichi [66]</b>	moderate	low	low	low
<b>Bobbio-Pallavicini [16]</b>	moderate	low	low	low
<b>Bos [17]</b>	moderate	low	low	low
<b>Baun-Moscovici [18]</b>	moderate	low	low	low
<b>Bienkowska [67]</b>	moderate	high	moderate	low
<b>Buch [19]</b>	moderate	low	low	low
<b>Cañete [20]</b>	low	moderate	high	low
<b>Chen [21]</b>	low	low	low	low
<b>Choi [22]</b>	moderate	moderate	low	low
<b>Couderc [23]</b>	low	low	low	low
<b>Coulthard [24]</b>	moderate	high	moderate	low
<b>Daïen [25]</b>	low	low	low	moderate
<b>Dávila-Fajardo [26]</b>	low	moderate	low	low
<b>Fabris [27]</b>	low	low	moderate	low
<b>Fabris [28]</b>	low	moderate	low	low
<b>Fisher [29]</b>	low	high	moderate	low
<b>Gottenberg [30]</b>	low	low	moderate	low
<b>Guis [31]</b>	low	low	low	low
<b>Hassan [32]</b>	moderate	high	low	low
<b>Hyrich [33]</b>	low	moderate	high	low
<b>Jamnitski [34]</b>	low	high	low	low
<b>Jancic [35]</b>	moderate	low	low	low
<b>Julia [68]</b>	low	low	moderate	low
<b>Kang [36]</b>	moderate	low	low	low
<b>Kastbom [37]</b>	low	low	low	low
<b>Kayakabe [38]</b>	low	low	low	low
<b>Klaasen [39]</b>	moderate	low	high	low
<b>Klaasen [40]</b>	moderate	low	high	low
<b>Klaasen [41]</b>	moderate	high	high	low
<b>Lequerré [42]</b>	moderate	low	moderate	low
<b>Lequerré [69]</b>	low	low	low	low
<b>Magnussen [43]</b>	moderate	low	low	low
<b>Marotte [44]</b>	moderate	low	low	low
<b>Matsudaira [45]</b>	moderate	low	low	low
<b>Morales-Lara [46]</b>	moderate	moderate	low	low
<b>Morales-Lara [47]</b>	moderate	low	low	low
<b>Mugnier [48]</b>	low	low	low	low
<b>Pers [49]</b>	high	moderate	low	low
<b>Pers [50]</b>	low	high	moderate	low
<b>Potter [51]</b>	low	high	low	low

---

<b>Quartuccio [52]</b>	low	moderate	low	low
<b>Quartuccio [53]</b>	low	moderate	moderate	low
<b>Radstake [54]</b>	moderate	moderate	low	low
<b>Rooryk [55]</b>	moderate	high	low	low
<b>Ruysen-Witrand [56]</b>	low	low	low	low
<b>Ruysen-Witrand [57]</b>	low	low	low	low
<b>Soliman [58]</b>	low	moderate	high	low
<b>Sellam [59]</b>	low	moderate	low	low
<b>Straub [60]</b>	moderate	low	moderate	low
<b>Stuhlmüller [61]</b>	low	moderate	low	low
<b>Tanino [70]</b>	moderate	low	moderate	low
<b>Thurlings [71]</b>	moderate	low	low	low
<b>Toonen [62]</b>	moderate	moderate	moderate	low
<b>Yukawa [64]</b>	low	high	moderate	moderate
<b>Wijbrandts [63]</b>	moderate	low	moderate	low

---

**Supplementary Table 2.** Multivariable biomarker models (without correction for clinical variables).

Author (et al.)	Class of biological	Predictors	Analyzed Cohort (n)	Prior prob per study	PPV (%)	NPV (%)	APV+	APV-
<b>Bansard † [65]</b>	ANA	7 RNA transcripts	18	44.4%	77.8%	88.9%	33.3%	-33.3%
<b>Bazzichi † [66]</b>	TNF	Early response calculator (ERI)	121	82.6%	84.6%	29.4%	2.0%	-12.1%
<b>Bienkowska †[67]</b>	TNF	7 RNA transcript	11	81.8%	90.0%	100.0%	8.2%	-81.8%
<b>Julia † [68]</b>	TNF	8 genes	14	85.7%	91.7%	50.0%	6.0%	-35.7%
<b>Lequerre [69]</b>	TNF	20 RNA transcript	20	50.0%	75.0%	87.5%	25.0%	-37.5%
<b>Lequerre [69]</b>	TNF	8 RNA transcript	20	50.0%	100.0%	83.3%	50.0%	-33.3%
<b>Tanino † [70]</b>	TNF	8 RNA transcripts	24	70.8%	80.0%	44.4%	9.2%	-15.3%
<b>Thurlings † [71]</b>	RTX	3 RNA transcripts	51	60.8%	70.4%	50.0%	9.6%	-10.8%

† = data extracted from validation cohort

Abbreviations:PPV=positive predictive value (i.e. the proportion of test positives which are responders (true positives) among all test positives), NPV=negative predictive value (i.e. the proportion of test negatives which are non-responders (true negatives) among all test negatives), APV=added predictive value (i.e. APV+/APV- are the resp. increase and decrease in absolute chance of response when the test-outcome (high or low value of the (bio)marker) is known.



## Chapter 3

### **Necessity of TNF-alpha inhibitor discontinuation in rheumatoid arthritis is predicted by smoking and number of previously used biological DMARDs**

Bart V.J. Cuppen, Johannes W.G. Jacobs, Evert-Jan ter Borg, Anne C.A. Marijnissen, Johannes W.J. Bijlsma, Floris P.J.G. Lafeber, Jacob M. van Laar, on behalf of all SRU investigators

## **ABSTRACT**

### **Objective**

Despite the success of TNF-alpha inhibitor (TNFi) treatment in rheumatoid arthritis (RA), a substantial number of patients necessitate discontinuation. Prediction thereof would be clinically relevant and guide the decision whether to start TNFi treatment.

### **Methods**

Data were used from the observational BiOCURA cohort, in which patients initiating biological treatment were enrolled and followed up for one year. In the model development cohort (n=192), a model predicting TNFi discontinuation was built using Cox-regression with backward selection ( $p < 0.05$ ). The parameters of the model were tested again in a model refinement cohort (n=60), for significance ( $p < 0.05$ ) and consistency of effect. In addition, we performed a systematic review to put our study results into perspective.

### **Results**

Of the 252 patients who initiated TNFi treatment, 103 (41%) had to discontinue treatment. Discontinuation was predicted at baseline by female gender, current smoking, high VAS-general health, and higher number of previously used biological disease-modifying anti-rheumatic drugs (bDMARDs). At refinement, smoking status and number of previously used bDMARDs remained with re-estimated hazard ratio's (HRs) in the total cohort of 1.74 (95%-CI 1.15-2.63,  $p < 0.01$ ) and 1.40 (95%-CI 1.16-1.68,  $p < 0.01$ ), respectively. Using these two predictors, we developed a simple score predicting discontinuation (PPV=72.3%). From literature, predictors were pack years of smoking, number of previously used bDMARDs, lack of any concomitant DMARD therapy and in particular lack of concomitant methotrexate (MTX).

### **Conclusion**

TNFi discontinuation is predicted by current smoking and number of previously used biological DMARDs, as well as by pack years of smoking and lack of any concomitant DMARD/MTX therapy.

## **INTRODUCTION**

Rheumatoid arthritis (RA) is a chronic, disabling immune disease predominantly involving synovial joints, affecting 0.5-1% of the population in industrialized countries [1-3]. Tumor necrosis factor-alpha inhibitor (TNFi) treatments have dramatically improved the outcome of RA patients. However, a substantial number of RA patients have an unsatisfactory response to these TNFi or experiences adverse-effects, necessitating discontinuation of therapy. Treatment success in clinical practice is a matter of balance between drug efficacy and tolerability; which continuously has to be weighed against the possibility of superior success to another treatment.

Prediction of discontinuation would give insight in patients at risk for treatment failure, and might indicate relevant factors for physician and patient when contemplating TNFi treatment. Exploratory studies have addressed possible predictors for discontinuation, although these have usually not been validated in a separate cohort. In addition, to our knowledge, no one has ever systemically analysed all reported predictors from literature to investigate whether these are consistent across studies. This abundance in unreplicated predictors that have never jointly been reviewed, leads to a vague idea of what possible predictors could look like in practice, though at the same time results in the absence of any certainties. The aim of the present study was therefore to find predictors for discontinuation and replicate these in a second cohort, and additionally, compare the predictors found in our study with those reported in literature by performing a systematic literature review.

## **METHODS**

### **Patients**

Data were used from the “Biologicals and Outcome Compared and predicted Utrecht region in Rheumatoid Arthritis” (BiOCURA) observational cohort, in which the first patient was enrolled in June 2009 and the last in March 2015. In BiOCURA, RA patients eligible for biological treatment according to regular clinical practice in 8 hospitals from the Society for Rheumatology research Utrecht (SRU, see Supplementary Text 1 for participating centres) in the Netherlands were followed up for one year after start of biological

treatment, or shorter in case of treatment discontinuation. Re-inclusion after switching to a different biological treatment was possible, at which patients entered again at baseline. We used data of all patients initiating TNFi (adalimumab, etanercept, infliximab, golimumab, and certolizumab pegol). To reflect clinical practice, no patients were excluded.

### **Data collection**

Trained research nurses collected all data at baseline and during follow-up, including patient history, disease activity assessments, concomitant treatment(s), serology status regarding rheumatoid factor (RF) or anti-citrullinated protein antibody (ACPA), body mass index (BMI), smoking status, alcohol use, several questionnaires, and drug survival. Necessity of discontinuation was defined as having to cease the initiated TNFi permanently within one year. Temporarily discontinuation because of certain circumstances (e.g. infections, surgery) and complete cessation in case of remission were not considered as necessity of discontinuation. Reasons for discontinuation were categorised as 'inefficacy', 'adverse events' (AEs, including systemic and local allergic reactions, infectious complications, newly diagnosed cancers, and other events that could be related to the treatment), 'combination of AEs and inefficacy', and 'other reasons'. Drug survival was calculated from the first dose of the TNFi until the last dose, for a maximum of one year. The BiOCURA study was approved by the ethical committee of the University Medical Center Utrecht and the institutional review boards of the participating centres. All patients provided written informed consent.

### **Systematic review**

We performed a systematic review of predictors for drug survival of TNFi therapy in RA to compare our findings with current knowledge on reported predictors for continuation or discontinuation. A search was performed on the 15<sup>th</sup> of July 2015 in PubMed and EMBASE using a search string including synonyms for RA, TNFi and (dis)continuation (see Supplementary Figure 1). After removal of duplicates, the articles were screened on title and abstract based on the inclusion criteria. Subsequently, articles were screened for the exclusion criteria after which studies for data-extraction remained (see Supplementary

Table 1). In order to summarise all studies evaluating the same parameters, we used the same direction of each predictor (e.g. methotrexate (MTX) use = “yes”) and determined if the predictor was significantly ( $p < 0.05$ ) associated with either continuation or discontinuation, or neither of these ( $p > 0.05$ ). Predictors were extracted from multivariable models if possible. Given conflicting results of studies, predictors for continuation or discontinuation were considered as “associated with TNFi drug survival” if they met the following two criteria: they had to have been investigated and reported in at least three original publications and they had to have been shown predictive for either continuation or discontinuation in more than a third of the respective studies.

### **Statistical analyses**

We assigned all patients initiating TNFi from June 2009 to October 2012 to a model development cohort ( $n=192$ ). All subsequent patients initiating TNFi until April 2015 with a year follow-up were assigned to a model refinement cohort ( $n=60$ ). Because we used for refinement this non-randomly selected cohort, the verification can be considered as in-between internal and external [4]. The rationale of this is that a non-random split contributes to differences in baseline characteristics between the model development cohort and the verification cohort, which increases the reliability of predictors if they nevertheless can be replicated. Independent sample t-tests and chi-square tests were used to test for statistical differences between the cohorts. Kaplan Meier curves with cox-regression were used to compare differences in survival between cohorts and subgroups for different reasons for discontinuation. A multiple imputation process (10 databases) was used to account for missing baseline data, which was present for a maximum of 15% per predictor. In the model development cohort, a univariate pre-selection ( $p < 0.2$ ) with cox-regression was performed to select predictive variables out of 24 clinical parameters (comorbidity, education level, C-reactive protein (CRP), EuroQol 5 dimensions (EQ5D), health assessment questionnaire (HAQ) and all listed in Table 1, except the individual TNFi’s as the choice is usually not at random). The parameters selected via this method were then included in a multivariable cox-regression model with backward selection ( $p < 0.05$ ) to build the development prediction model. This model was re-applied in the model refinement cohort, in which the selected variables were re-tested on significance

**Table 1:** Baseline characteristics of model development and model refinement cohort. All patients in BiOCURA initiating TNFi from June 2009 to October 2012 were assigned to the model development cohort (n=192), and all subsequent patients initiating TNFi until April 2015 to the model refinement cohort (n=60). The presented clinical characteristics are all before treatment initiation. P-values for comparisons were calculated by means of an independent sample t-test, Mann-Whitney U test or chi-square based on distribution of the clinical parameter.

	<b>Model development cohort (n=192)</b>	<b>Model refinement cohort (n=60)</b>	<b>P-value</b>
<b>Gender, female, n (%)</b>	145 (75.5)	43 (71.7)	0.55
<b>Age, mean (SD)</b>	54.0 (±12.3)	57.1 (±10.7)	0.06
<b>Disease duration, median (IQR)</b>	5.0 (2.0-11.8)	6.0 (2.0-15.8)	0.69
<b>Smoking status, current, n (%)</b>	46 (24.0)	21 (35.0)	0.09
<b>Alcohol use, &gt;7 units/week, n (%)</b>	32 (16.8)	10 (16.7)	0.99
<b>BMI, mean (SD)</b>	27.0 (±5.2)	27.6 (±5.9)	0.48
<b>RF, positive, n (%)</b>	125 (65.8)	34 (56.7)	0.20
<b>ACPA, positive, n (%)</b>	134 (71.3)	41 (68.3)	0.66
<b>Baseline DAS28, mean (sd)</b>	4.5 (±1.2)	4.0 (±1.2)	0.02
<b>TJC, median (IQR)</b>	6 (2-12)	5 (2-11)	0.40
<b>SJC, median (IQR)</b>	2 (0*-4)	1 (0*-2)	<0.01
<b>ESR, median (IQR)</b>	19.0 (10-34)	14 (7-34)	0.15
<b>VAS-GH, mean (SD)</b>	55.9 (±24.0)	50 (±21.5)	0.09
<b>No. of previously used bDMARDs, n (%)</b>			0.85
<b>0 (naïve)</b>	119 (62.0)	33 (55.0)	
<b>1</b>	59 (30.7)	20 (33.3)	
<b>2</b>	8 (4.2)	4 (6.7)	
<b>&gt;2</b>	6 (3.1)	3 (5.0)	
<b>Initiated TNFi, n (%)</b>	192 (100.0)	60 (100.0)	0.20
<b>Adalimumab</b>	74 (38.5)	17 (28.3)	
<b>Etanercept</b>	68 (35.4)	29 (48.3)	
<b>Golimumab</b>	28 (14.6)	5 (8.3)	
<b>Infliximab</b>	11 (5.7)	3 (5.0)	
<b>Certolizumab</b>	11 (5.7)	6 (10.0)	
<b>Concomitant MTX, n (%)</b>	137 (71.4)	46 (76.7)	0.42
<b>Concomitant HCQ, n (%)</b>	51 (26.6)	15 (25.0)	0.81
<b>Concomitant SSZ, n (%)</b>	18 (9.4)	4 (6.7)	0.52
<b>Concomitant GC, n (%)</b>	69 (35.9)	21 (35.0)	0.90

\*values for swollen joint count of '0' were usually seen in patients who switched due to side effect and/or had involvement of joints outside 28-joint count.

ACPA, anti-citrullinated protein antibody; bDMARDs, biological disease modifying anti-rheumatic drugs; BMI, body mass index; DAS28, disease activity score based on 28 joints; GC, glucocorticoid; HCQ, hydroxychloroquine; MTX, methotrexate; RF, rheumatoid factor; SSZ, sulfasalazine; SJC, swollen joint count; TJC, tender joint count; TNFi, tumor necrosis factor alpha inhibitor; VAS-GH, visual analogue scale general health.

( $p < 0.05$ ) and consistency of effect (same direction of coefficient), and excluded if either of the two was violated. To develop a simple score for daily clinical practice distinguishing between patients with high and general risk for necessity of discontinuation, the final model was re-applied in the total cohort ( $n=252$ ). This yielded decimal regression coefficients, not easy to use for a simple prediction score; we therefore converted the coefficients by multiplying these into (nearly) integers. Risk scores for several cut-offs were compared to find a positive predictive value (PPV) for discontinuation of at least 70%. All analyses were performed using SPSS (v. 21.0. Released 2012. IBM Corp. Armonk, NY, USA).

## RESULTS

Baseline characteristics for patients of the model development and refinement cohorts are shown in Table 1. Differences in characteristics were seen due to a non-random split of the cohort, with most importantly a lower baseline DAS28 (mean 4.0 *versus* 4.5,  $p=0.02$ ), including a lower swollen joint count (median of 1 *versus* 2,  $p < 0.01$ ) in the model refinement cohort, compared to model development cohort. Within 1 year of follow-up 103 of the 252 patients (41%) discontinued TNFi treatment, with an incidence rate of 55 per 100 patient years. Patients most frequently discontinued between 10 and 26 weeks after start of treatment (Figure 1a). The drug survival was not significantly different between the model development and refinement cohort ( $p=0.38$ ). Inefficacy was the most common reason for discontinuation (62%), followed by AEs (31%) and AEs with inefficacy (6%) (Table 2). The least favourable TNFi survival was seen in patients discontinuing because of both AEs and inefficacy (Figure 1b).

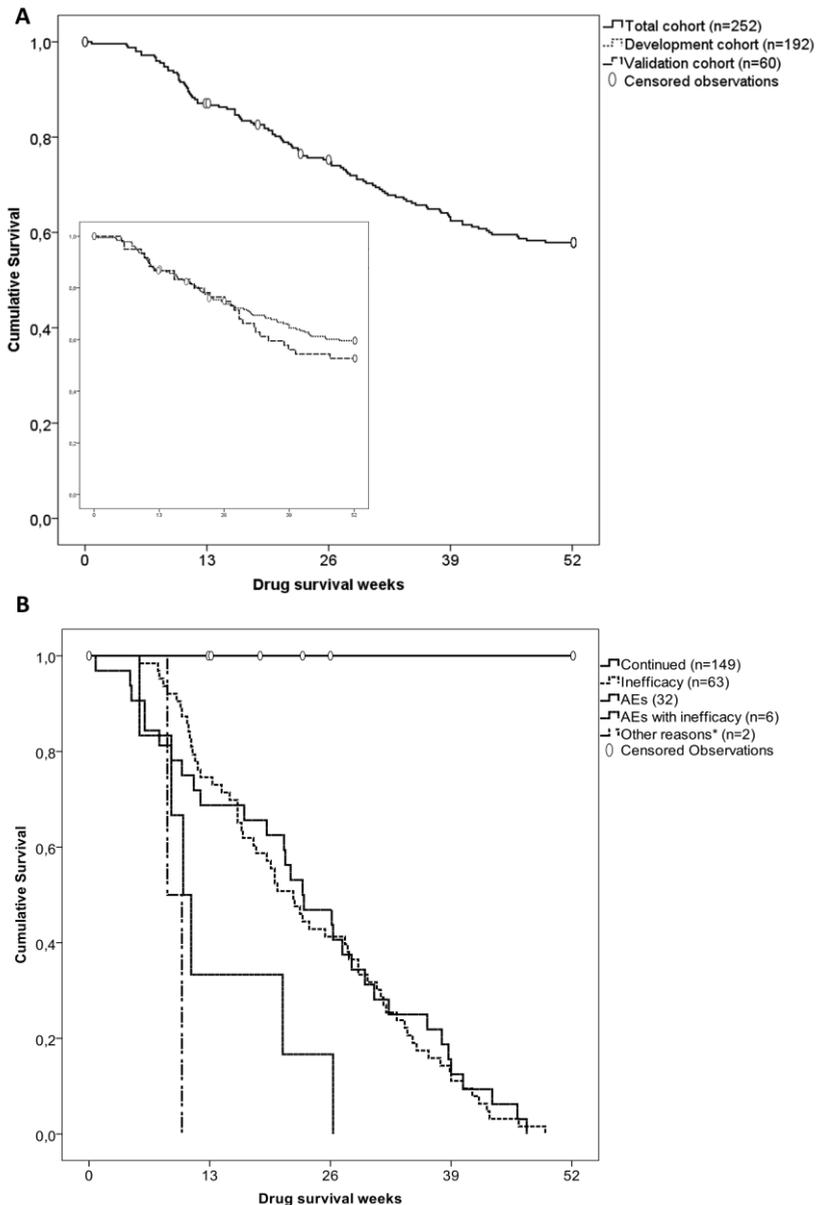
**Table 2:** Reasons for discontinuation within the first year after TNFi initiation. Shown are the number of discontinuing patients in model development, model refinement and complete cohort.

<b>Reason for discontinuation</b>	<b>Model development cohort (n=192)</b>	<b>Model refinement cohort (n=60)</b>	<b>Total cohort (n=252)</b>
<b>Inefficacy, n (% of total discontinuations)†</b>	48 (64)	15 (54)	63 (61)
<b>AEs, n (% of total discontinuations)</b>	21 (28)	11 (39)	32 (31)
<b>AEs with inefficacy, n (% of total discontinuations)</b>	4 (5)	2 (7)	6 (6)
<b>Other reasons*, n (% of total discontinuations)</b>	2 (3)	0 (0)	2 (2)
<b>Total, n (% of total discontinuations)</b>	<b>75 (100)</b>	<b>28 (100)</b>	<b>103 (100)</b>

\*1 patient discontinued because she had a pregnancy wish, 1 patient discontinued for unknown reasons.

† Of all 48 patients discontinuing due to inefficacy in the development cohort, n=37 (77.1%) had a primary failure (i.e. never experienced a clinical response at all), and n=11 (22.9%) had a secondary failure (i.e. lost their initial beneficial response to TNFi treatment). Of all 15 patients discontinuing due to inefficacy in the validation cohort, n=10 (66.7%) had a primary failure and n=5 (33.3%) a secondary failure.

AEs, adverse events; TNFi, tumor necrosis factor alpha inhibitor.



**Figure 1:** TNFi drug survival curves. **A:** TNFi drug survival curve for all patients (n=252) and separately for model development (n=192) and model refinement cohorts (n=60). There was no difference in survival between the model development and model refinement cohort ( $p=0.38$ , cox -regression, corrected for age and gender). **B:** TNFi drug survival curves for patients discontinuing for specified reasons. There was no difference in survival between patients discontinuing because of inefficacy and adverse events (AEs) without or with inefficacy ( $p=0.74$ , cox-regression after combining “AEs with inefficacy” with “AEs” to one group (n=38) and compare with inefficacy (n=63), corrected for age and gender).

In the model development cohort, nine of the 24 baseline clinical parameters were associated with discontinuation. These parameters were [HR (95%-confidence interval), p-value]: female gender [1.71 (0.94-3.12), p=0.08], current smoking (as opposed to only smoked in the past and never smoked) [1.45 (0.88-2.38), p=0.14], rheumatoid factor positivity [0.72 (0.45-1.15), p=0.17], concomitant use of MTX [0.61 (0.38-0.99), p=0.04], number of previously used biological DMARDs (bDMARDs) [1.42 (1.13-1.77) per biological, p<0.01], baseline DAS28 [1.27 (1.05-1.54), p=0.01], visual analogue scale of general health [VAS-GH, 1.21 (1.09-1.34)/per 10 mm, p<0.01], EQ5D [0.32 (0.15-0.69), p<0.01], and the HAQ, [1.38 (0.95-2.00), p=0.09]. The multivariable prediction model after backward selection is shown in Table 3. Discontinuation was predicted by a model including female gender [2.06 (1.12-3.80), p=0.02], current smoking [1.67 (1.00-2.77), p=0.05], higher number of previously used bDMARDs [1.46 (1.15-1.86) per biological, p<0.01], and high VAS-GH score [1.20 (1.08-1.33) per 10mm, p<0.01]. These four clinical parameters were re-tested in the model refinement cohort for significance and consistency of effects, resulting in exclusion of VAS-GH score [0.98 (0.84-1.15), p=0.90] and female gender [0.84 (0.34-2.06), p=0.70]. The model that remained contained the predictors current smoking [2.53 (1.14-5.61), p=0.02] and number of previously used bDMARDs [1.43 (1.03-1.99) per biological, p=0.03].

**Table 3:** Multivariable Cox model predicting risk of discontinuation. The model was built by entering the 9 predictors from univariate pre-selection and performing backward selection (p<0.05) on all patients in the model development cohort (n=192). The categorical variable smoking (current/only in the past/never) was dichotomised in current smoking (yes/no) based on the results of univariate analysis, only showing an effect for current smoking.

<b>Development</b>	<b>HR</b>	<b>95%-CI</b>	<b>P-value</b>
<b>Female Gender, yes</b>	2.06	(1.12-3.80)	0.02
<b>Current smoking, yes</b>	1.66	(1.00-2.77)	0.05
<b>No. of prev. biologicals, per biological</b>	1.46	(1.15-1.86)	<0.01
<b>VAS-GH, per 10mm</b>	1.20	(1.08-1.33)	<0.01

CI, confidence interval; HR, hazard ratio; VAS-GH, visual analogue scale general health (0-100, 0=best health).

To create a simple prediction score of use for clinical practice, we first re-estimated the regression coefficients by applying the final model to the complete cohort, which resulted in a HR of 1.74 for current smoking and HR of 1.40 per number of previously used bDMARDs (see Table 4). The subsequently constructed risk score is used by assigning points per patient for the two predictors (5 for smokers, 3 per previously used bDMARD) and summarizes these points. For possible cut-offs, the predicted chances on discontinuation were evaluated (see Supplementary Table 2). The cut-off of  $\geq 9$  points predicted an absolute risk of discontinuation (PPV) of 72.3%, and was scored by 11 patients (4.4%) of the total study cohort. This cut-off is met by non-smokers with three or more-, and smokers with two or more previously used bDMARDs.

**Table 4:** Final prediction model and simple prediction score for use in daily practice. After refinement, the remaining parameters were current smoking and number of previously used bDMARDs. Shown are the estimated predictive values of these remaining parameters in the total cohort (n=252). The regression coefficients of smoking (0.553) and number of previously used biologicals (0.334) were multiplied by 9 to create nearly integers to use as assigned points (5 and 3 respectively). Scores assigned to these predictors have to be summed to get the total score. The cut-off of  $\geq 9$  met our criteria of a positive predictive value  $>0.7$ , which is scored by 4.4% of the complete cohort. This cut-off is met by non-smokers with 3 or more- and smokers with 2 or more previously used bDMARDs, who have a high risk of having to discontinue the initiated TNFi.

	HR	95%-CI of HR	P-value	Assigned points	Score
<b>Current smoking, yes</b>	1.74	(1.15-2.63)	<0.01	5	....
<b>No. of prev. used biologicals, per biological</b>	1.40	(1.16-1.68)	<0.01	3	....
<b>Total Score</b>					....

CI, confidence interval; HR, hazard ratio.

To put our results in context we performed a systematic literature search. The search resulted in 3311 articles, of which 92 on predicting of continuation or discontinuation remained after screening (see Supplementary Figure 1). After applying the exclusion criteria, 30 full-text articles and 18 congress abstracts remained. Results for the full-text articles and predictors that were studied in  $\geq 3$  studies are shown in Table 5, whereas all

investigated predictors and the results including (congress) abstracts are shown in Supplementary Table 3.

**Table 5:** Results from systematic review on investigated predictors for TNFi drug survival, reported in three or more studies. Predictors were drawn from multivariable models if possible and a p-value <0.05 was considered an effect. A predictor from a study was assigned to either one of three categories: associated with continuation, discontinuation or not associated. Per category, the number of studies that mention the particular predictor is shown. The association with survival (i.e. conclusion) is based on the consistency of the predictor in predicting the ‘same direction’ in more than 1/3rd of the studies. The references of all studies per predictor can be found in supplementary table 3.

Item	Association with continuation	No association with survival	Association with discontinuation	Conclusion TNFi drug survival
<b>Demographic and clinical parameters</b>				
Age*	0	16	3	-
Female gender	1	16	1	-
Disease-duration	2	10	0	-
Comorbidity §	1	7	1	-
RF positivity/levels	0	7	0	-
Current smoking	0	2	1	-
Pack years smoking	0	1	2	†
<b>Assessments</b>				
DAS28*	1	11	3	-
HAQ	0	8	3	-
SJC	0	6	0	-
ESR	1	4	1	-
TJC	0	4	2	-
CRP	1	5	0	-
<b>Treatments</b>				
Concomitant MTX	6	6	2	‡
Concomitant GCs	1	8	1	-
Number of previously used DMARDs*	0	6	2	-
Any concomitant csDMARD	5	0	1	‡
Number previously used bDMARDs*§	0	1	4	†
Concomitant SSZ	0	2	1	-

\* combination of numerical and categorical values for this parameter

§ different categories of parameters per study investigated

‡ predicting continuation (associated in >1/3rd of studies with continuation)

- no association with TNFi drug survival

† predicting discontinuation (associated in >1/3rd of studies with discontinuation)

ANA, anti-nuclear antibodies; bDMARD, biological DMARD; csDMARD, conventional synthetical DMARD; DAS28, disease activity score based on 28 joints; DMARDs, disease modifying antirheumatic drugs; ESR, erythrocyte sedimentation rate; GCs, glucocorticoids; HAQ, health assessment questionnaire; MTX, methotrexate; NSAID, nonsteroidal anti-inflammatory drugs; RF, rheumatoid factor; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count.

The predictors significantly associated with continuation were concomitant MTX use and any concomitant DMARD use, whereas those associated with discontinuation were number of previously used bDMARDs and pack years smoking. In addition, smoking status was significantly associated with discontinuation in one study and borderline-significant in two other studies [5, 6].

## **DISCUSSION**

Our study shows that TNFi discontinuation can be predicted by smoking status and number of previously used bDMARDs, which is supported by results from other observational studies reported in literature. A simple prediction score was able to predict TNFi discontinuation with accuracy (PPV=72.3%), although was only applicable to a minority of all patients (i.e. 4.4%). According to this prediction score, smokers with a history of  $\geq 2$  previously used biologicals, and non-smokers with  $\geq 3$  biologicals represent a group of RA patients with a low TNFi treatment success. The systematic review also revealed pack years of smoking, lack of any concomitant DMARD therapy and lack of concomitant MTX as predictors for TNFi discontinuation.

Smoking status was investigated in three previous studies [5-7]. Although a significant association with TNFi discontinuation was only found in one of these studies [7], the two other studies reported trends in the same direction, with p-values of 0.055[5] and 0.075 [6]. In addition, when our results would be added to the systematic review, current smoking would fulfill the criteria (significant in 2/4 studies). Also of note, in one abstract covering data of 12,000 TNFi users, an odds ratio for smoking of 1.20 (1.06-1.36) was found [8]. Our results, together with the results from literature therefore indicate that smoking status is a rather moderate, but robust predictor for discontinuation. Smoking is known to increase susceptibility of RA [9-11], is associated with a higher disease activity [12-16] and a reduced clinical response to MTX and TNFi therapy [17-20]. Several mechanisms could play a role in the reduced clinical response to treatment. First, in smoking RA patients ACPA levels are higher [14, 17, 21-25], and higher levels are associated with a more severe disease course [26-31]. However, when the relationship

between smoking and response to therapy is tested, the reduced clinical response seems to be independent from ACPA levels [20]; similarly, ACPA positivity was not selected at our univariate pre-selection, but smoking was. These data indicate that smoking is a more robust predictor for discontinuation than ACPA. Second, various serum cytokines and matrix metalloproteinases levels are elevated in smokers which could make smoking patients more resistant to TNFi therapy [17, 32, 33]. Third, several pharmaco-kinetic or –dynamic effects could play a role [19, 20]. For example, both current smoking and systemic inflammation elevate basal metabolic rate [34], and may reduce the bioavailability of anti-rheumatic drugs in smokers in comparison with non-smokers [24]. Fourth, smoking status may also be an indicator of lower socioeconomic status, which is related with a poorer outcome of RA [35]. Regardless of the underlying mechanism, it is difficult to use smoking status in clinical decision making, as it seems to negatively affect responsiveness to a range of treatments [17-20], limiting the choice of effective alternatives, although no negative effect of smoking on the efficacy of tocilizumab was found [36]. In addition, it has not been proven after quitting of smoking, the effects of TNFi's will be better. Nevertheless, because the risk of cardiovascular morbidity and mortality in RA is already increased, even in non-smoking patients [37], smokers should always be encouraged to quit.

The number of previously used biologicals showed a HR of 1.40 (1.16-1.68) in the BiOCURA cohort, and has been reported as a predictor for discontinuation in four out of five studies [38-41] and in all three additional abstracts [42-44]. Selection may be the main cause for the predictive ability of this parameter: among non-naïve TNFi patients probably a higher number of TNFi treatment refractory patients was found. It is tempting to speculate that primary refractory patients to TNFi treatment have a type of RA that is probably more dependent on other inflammatory mechanisms than the TNF-alpha pathway, which increases the chance for a second failure to TNFi treatment [38]. However, although this makes sense, observations in clinical practice do not support this hypothesis [45]. Another explanation for a TNFi refractory RA might be the development of anti-drug antibodies that eliminate the TNFi until subclinical concentrations are reached, leading to a non-response or loss of response. It has been found that patients who developed antibodies to

infliximab therapy, are more prone to develop anti-drug antibodies to the next TNFi (adalimumab), compared with infliximab naïve patient that initiate adalimumab [46], which could explain why patients failing a previous TNFi have reduced chances for the next TNFi. However, when all switchers from infliximab to adalimumab were investigated, response to adalimumab was higher in switchers with detectable anti-infliximab antibodies than in switchers without these antibodies [46, 47], which challenges the hypothesis above.

Lack of any concomitant DMARD and in particular lack of MTX predicted discontinuation of TNFi therapy in literature. In our analysis, concomitant MTX was protective for discontinuation in univariate analysis (HR=0.61), though it was excluded in multivariable analysis. It has been shown that MTX reduces the immunogenicity to TNFi [48]. Therefore, even when response to MTX monotherapy is insufficient, continuation of MTX is advised when commencing a bDMARD [49]. Female gender and a higher VAS-GH were found to be associated with discontinuation in our model development cohort, however, these predictors could not be replicated in our model refinement cohort. The inability to replicate is in line with the results from literature, which also showed no true association between these predictors and TNFi drug survival (Supplementary Table 3).

The limited number of included patients may be considered a limitation of our study. However, our findings are consistent with the results from literature, which reduces the possibility false positive results were shown (type I error). We cannot exclude the possibility though of any predictors not found due to a low power, i.e. false negative results (type II error). However, any factor that we have missed due to a low power, would probably have a lesser effect on discontinuation than the predictors we did find, which therefore have less clinical relevance. Also, of all included patients, 31 were re-included for a second treatment course (i.e. entered the study twice). In order to analyse if this influenced the results, as a sensitivity analysis the final model was re-applied on the total cohort while excluding all re-inclusions. Only changes in predictive values <10% were observed for smoking status and number of previously used biologicals, demonstrating absence of major bias due to re-inclusion. It should also be noted that model development

techniques such as backward selection, increase the chance of model overfitting. It might therefore be possible that the true HRs of the predictors are lower than estimated, or even that any true association with discontinuation is absent. However, the fact that the predictors we found could be validated in literature, makes it very unlikely these are false positive results. The systematic literature search was carried out using a broad search string and included many studies and several abstracts. However, the heterogeneity among studies was not taken into account, which might have influenced the categorisation of several factors. The presented data should therefore be regarded as a concise, rather than an exhaustive overview. Also, the limited predictive value in our study and other observational studies, could be a result of the subjectivity of the outcome (no strict criteria for discontinuation in practice) and the heterogeneity in the interaction between patients and physicians in clinical decision-making. Moreover, discontinuation is a multi-dimensional endpoint with many influential parameters, which inhibits the creation of a fully explanatory model containing only several (clinical) parameters. Although future research could add to the identification of better predictors, a relatively wide error margin seems inevitable when using discontinuation as an outcome.

## **CONCLUSION**

We showed that necessity of TNFi discontinuation is related to smoking status and number of previously used biologicals, which is corroborated by our systemic literature search and also revealed pack years of smoking, lack of any concomitant DMARD therapy and lack of concomitant MTX as predictors. The exact mechanisms of action leading to discontinuation are not always known, and are complicated by the multidimensional complexity that leads to the clinical decision of discontinuation.

## REFERENCES

1. Wolfe AM, Kellgren JH, Masi AT. The epidemiology of rheumatoid arthritis: a review. II. Incidence and diagnostic criteria. *Bull Rheum Dis* 1968;19 3: 524-9.
2. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res* 2002;4 Suppl 3: S265-S72.
3. Gabriel SE, Michaud K. Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. *Arthritis Res Ther* 2009;11 3: 229.
4. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162 1: W1-73.
5. Figueiredo IT, Morel J, Sany J, Combe B. Maintenance and tolerability of infliximab in a cohort of 152 patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2008;26 1: 18-23.
6. Soderlin MK, Petersson IF, Geborek P. The effect of smoking on response and drug survival in rheumatoid arthritis patients treated with their first anti-TNF drug. *Scand J Rheumatol* 2012;41 1: 1-9.
7. Matthey DL, Dawes PT, Hassell AB, Brownfield A, Packham JC. Effect of psychological distress on continuation of anti-tumor necrosis factor therapy in patients with rheumatoid arthritis. *J Rheumatol* 2010;37 10: 2021-4.
8. McWilliams DF, Walsh DA. Factors predicting pain and early discontinuation of tumour necrosis factor-alpha-inhibitors in people with rheumatoid arthritis: Results from the British society for rheumatology biologics register. *Ann Rheum Dis* 2014;73 SUPPL. 2: 602.
9. Kallberg H, Padyukov L, Plenge RM, Ronnelid J, Gregersen PK, AH vdH-vM et al. Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. *Am J Hum Genet* 2007;80 5: 867-75.
10. Pedersen M, Jacobsen S, Garred P, Madsen HO, Klarlund M, Svejgaard A et al. Strong combined gene-environment effects in anti-cyclic citrullinated peptide-positive rheumatoid arthritis: a nationwide case-control study in Denmark. *Arthritis Rheum* 2007;56 5: 1446-53.
11. Kokkonen H, Johansson M, Innala L, Jidell E, Rantapaa-Dahlqvist S. The PTPN22 1858C/T polymorphism is associated with anti-cyclic citrullinated peptide antibody-positive early rheumatoid arthritis in northern Sweden. *Arthritis Res Ther* 2007;9 3: R56.
12. Tureson C, Schaid DJ, Weyand CM, Jacobsson LT, Goronzy JJ, Petersson IF et al. Association of HLA-C3 and smoking with vasculitis in patients with rheumatoid arthritis. *Arthritis Rheum* 2006;54 9: 2776-83.
13. Manfredsdottir VF, Vikingsdottir T, Jonsson T, Geirsson AJ, Kjartansson O, Heimisdottir M et al. The effects of tobacco smoking and rheumatoid factor seropositivity on disease activity and joint damage in early rheumatoid arthritis. *Rheumatology (Oxford)* 2006;45 6: 734-40.
14. Wolfe F. The effect of smoking on clinical, laboratory, and radiographic status in rheumatoid arthritis. *J Rheumatol* 2000;27 3: 630-7.
15. Papadopoulos NG, Alamanos Y, Voulgari PV, Epagelis EK, Tsifetaki N, Drosos AA. Does cigarette smoking influence disease expression, activity and severity in early rheumatoid arthritis patients? *Clin Exp Rheumatol* 2005;23 6: 861-6.
16. Nyhall-Wahlin BM, Jacobsson LT, Petersson IF, Tureson C. Smoking is a strong risk factor for rheumatoid nodules in early rheumatoid arthritis. *Ann Rheum Dis* 2006;65 5: 601-6.
17. Matthey DL, Brownfield A, Dawes PT. Relationship between pack-year history of smoking and response to tumor necrosis factor antagonists in patients with rheumatoid arthritis. *J Rheumatol*. 2009;1180-7.
18. Hyrich KL, Watson KD, Silman AJ, Symmons DP. 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: 1558-65.
19. Abhishek A, Butt S, Gadsby K, Zhang W, Deighton CM. Anti-TNF-(alpha) agents are less effective for the treatment of rheumatoid arthritis in current smokers. *J Clin Rheumatol* 2010;16 1: 15-8.
20. Saevarsdottir S, Wedren S, Seddighzadeh M, Bengtsson C, Wesley A, Lindblad S et al. Patients with early rheumatoid arthritis who smoke are less likely to respond to treatment with methotrexate and tumor necrosis factor inhibitors: observations from the Epidemiological Investigation of Rheumatoid Arthritis and the Swedish Rheumatology Register cohorts. *Arthritis Rheum* 2011;63 1: 26-36.
21. Saag KG, Cerhan JR, Kolluri S, Ohashi K, Hunninghake GW, Schwartz DA. Cigarette

- smoking and rheumatoid arthritis severity. *Ann Rheum Dis* 1997;56 8: 463-9.
22. Padyukov L, Silva C, Stolt P, Alfredsson L, Klareskog L. A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. *Arthritis Rheum* 2004;50 10: 3085-92.
  23. Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum* 2006;54 1: 38-46.
  24. Westhoff G, Rau R, Zink A. Rheumatoid arthritis patients who smoke have a higher need for DMARDs and feel worse, but they do not have more joint damage than non-smokers of the same serological group. *Rheumatology (Oxford)* 2008;47 6: 849-54.
  25. Linn-Rasker SP, AH vdH-vM, van Gaalen FA, Kloppenburg M, de Vries RR, le CS et al. Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients who carry HLA-DRB1 shared epitope alleles. *Ann Rheum Dis* 2006;65 3: 366-71.
  26. van der Helm-van Mil A, Verpoort KN, Breedveld FC, Toes RE, Huizinga TW. Antibodies to citrullinated proteins and differences in clinical progression of rheumatoid arthritis. *Arthritis Res Ther* 2005;7 5: R949-R58.
  27. Sakkas LI, Bogdanos DP, Katsiari C, Platsoucas CD. Anti-citrullinated peptides as autoantigens in rheumatoid arthritis-relevance to treatment. *Autoimmun Rev* 2014;13 11: 1114-20.
  28. Kroot EJ, de Jong BA, van Leeuwen MA, Swinkels H, van den Hoogen FH, van't Hof M et al. The prognostic value of anti-cyclic citrullinated peptide antibody in patients with recent-onset rheumatoid arthritis. *Arthritis Rheum* 2000;43 8: 1831-5.
  29. Nell VP, Machold KP, Stamm TA, Eberl G, Heinzl H, Uffmann M et al. Autoantibody profiling as early diagnostic and prognostic tool for rheumatoid arthritis. *Ann Rheum Dis* 2005;64 12: 1731-6.
  30. Jorgensen C, Legouffe MC, Bologna C, Brochier J, Sany J. IgA isotype rheumatoid factor in rheumatoid arthritis: clinical implications. *Clin Exp Rheumatol* 1996;14 3: 301-4.
  31. Jonsson T, Arinbjarnarson S, Thorsteinsson J, Steinsson K, Geirsson AJ, Jonsson H et al. Raised IgA rheumatoid factor (RF) but not IgM RF or IgG RF is associated with extra-articular manifestations in rheumatoid arthritis. *Scand J Rheumatol* 1995;24 6: 372-5.
  32. Cozen W, Diaz-Sanchez D, James GW, Zadnick J, Cockburn MG, Gill PS et al. Th1 and Th2 cytokines and IgE levels in identical twins with varying levels of cigarette consumption. *J Clin Immunol* 2004;24 6: 617-22.
  33. Raitio A, Tuomas H, Kokkonen N, Salo T, Sorsa T, Hanemaaijer R et al. Levels of matrix metalloproteinase-2, -9 and -8 in the skin, serum and saliva of smokers and non-smokers. *Arch Dermatol Res* 2005;297 6: 242-8.
  34. Metsios GS, Stavropoulos-Kalinoglou A, Nevill AM, Douglas KM, Koutedakis Y, Kitas GD. Cigarette smoking significantly increases basal metabolic rate in patients with rheumatoid arthritis. *Ann Rheum Dis* 2008;67 1: 70-3.
  35. Harrison MJ, Farragher TM, Clarke AM, Manning SC, Bunn DK, Symmons DP. Association of functional outcome with both personal- and area-level socioeconomic inequalities in patients with inflammatory polyarthritis. *Arthritis Rheum* 2009;61 10: 1297-304.
  36. Pers YM, Fortunet C, Constant E, Lambert J, Godfrin-Valnet M, De JA et al. Predictors of response and remission in a large cohort of rheumatoid arthritis patients treated with tocilizumab in clinical practice. *Rheumatology (Oxford)* 2014;53 1: 76-84.
  37. Nurmohamed MT, Heslinga M, Kitas GD. Cardiovascular comorbidity in rheumatic diseases. *Nat Rev Rheumatol* 2015;11 12: 693-704.
  38. Agarwal SK, Glass RJ, Shadick NA, Coblyn JS, Anderson RJ, Maher NE et al. Predictors of discontinuation of tumor necrosis factor inhibitors in patients with rheumatoid arthritis. *J Rheumatol* 2008;35 9: 1737-44.
  39. Chatzidionysiou K, Kristensen LE, Eriksson J, Askling J, van Vollenhoven R, Group FTA. Effectiveness and survival-on-drug of certolizumab pegol in rheumatoid arthritis in clinical practice: results from the national Swedish register. *Scand J Rheumatol* 2015;18: 1-7.
  40. Finckh A, Simard JF, Gabay C, Guerne PA. Evidence for differential acquired drug resistance to anti-tumour necrosis factor agents in rheumatoid arthritis. *Ann Rheum Dis* 2006;65 6: 746-52.
  41. Martinez-Santana V, Gonzalez-Sarmiento E, Calleja-Hernandez M, Sanchez-Sanchez T. Comparison of drug survival rates for tumor necrosis factor antagonists in rheumatoid arthritis. *Patient Prefer Adherence* 2013;7: 719-27.

42. Gibofsky A, Saunders KC, Ganguli A, Cifaldi M, Grant S, Clewell J et al. Persistency and predictors of persistency of adalimumab among rheumatoid arthritis (RA) patients in a us registry. *Arthritis Rheum* 2012;64 SUPPL. 10: S165-S6.
43. Pocock JM, Albreshni S, Slack R, Polydoropoulou V, O'Reilly D, Ostor AJ. A comparison of drug survival times for infliximab, adalimumab and etanercept. *Rheumatology (Oxford)* 2009;48 SUPPL. 1: i69.
44. Saevarsdottir S, Santacatterina M, Turesson C, Forsblad H, Jacobsson L, Lindblad S. Clinical characteristics and outcome of golimumab treatment differs between bio-naive and patients previously exposed to biologicals. Nationwide results on Rheumatoid Arthritis (RA), Psoriatic Arthritis (PSA), ankylosing spondylitis (AS) and Other Spondylarthritidies (SPA). *Arthritis Rheum* 2014;66 SUPPL. 10: S705.
45. Smolen JS, Aletaha D. Forget personalised medicine and focus on abating disease activity. *Ann Rheum Dis* 2013;72 1: 3-6.
46. Bartelds GM, Wijbrandts CA, Nurmohamed MT, Stapel S, Lems WF, Aarden L et al. Anti-infliximab and anti-adalimumab antibodies in relation to response to adalimumab in infliximab switchers and anti-tumour necrosis factor naive patients: a cohort study. *Ann Rheum Dis* 2010;69 5: 817-21.
47. Jamnitski A, Bartelds GM, Nurmohamed MT, van Schouwenburg PA, van SD, Stapel SO et al. The presence or absence of antibodies to infliximab or adalimumab determines the outcome of switching to etanercept. *Ann Rheum Dis* 2011;70 2: 284-8.
48. Garces S, Demengeot J, Benito-Garcia E. The immunogenicity of anti-TNF therapy in immune-mediated inflammatory diseases: a systematic review of the literature with a meta-analysis. *Ann Rheum Dis* 2013;72 12: 1947-55.
49. Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73 3: 492-509

**Supplementary Text 1.** Society for Rheumatology Research Utrecht (SRU) contributing hospitals' contact persons:

dr. J. Tekstra, University Medical Center Utrecht  
dr. E.J. ter Borg Antonius, Hospital Nieuwegein/Utrecht  
drs. Y. Schenk, Diakonessen Hospital Utrecht  
dr. S. Linn-Rasker, Meander Medical Center Amersfoort  
dr. W.H. van der Laan, Sint Maartenskliniek Woerden  
drs. D.G. Kuiper-Geertsma, Hospital St. Jansdal Harderwijk  
dr. M. Nabibux, Tergooi Hospital Hilversum  
dr. C.M. Verhoef, Flevo Hospital Almere

**Supplementary Table 1:** Search syntax for PUBMED and EMBASE. The search was performed on title/abstract. The search string was designed according to the patients, intervention, comparison and outcome (PICO) method. No “comparison” was integrated in the search, since not all studies mention common descriptions such as “predictor”, but rather the predictors found.

Patients: (rheumatoid AND arthritis) OR RA

---

Intervention: (anti AND tnfalphabeta) OR anti-TNF OR antitnf OR (TNF AND antagonist) OR (TNF AND antagonists) OR (TNF AND inhibitor) OR (TNF AND inhibitors) OR Adalimumab OR Etanercept OR Infliximab OR Golimumab OR Certolizumab

---

Outcome: discontinuation OR discontinuing OR continuation OR continuing OR survival OR stopping OR retention OR attrition OR (treatment AND failure) OR (treatment AND success)

**Supplementary Table 2:** Cut-offs for the simple prediction score and absolute chances on discontinuation. A cut-off of 9 met our predefined criterion of a PPV  $\geq$  70% and is able to classify 11/252 (=4.4%) of patients as high risk.

<b>Used cut-off for total prediction score</b>	<b>No. of patients defined as high risk by cut-off, n (%)</b>	<b>No. of patients discontinuing within high risk patients, n (%)</b>
<b>No cut-off</b>	252 (100)	103 (41)
<b><math>\geq 3</math></b>	141 (56)	71 (50)
<b><math>\geq 5</math></b>	86 (34)	47 (55)
<b><math>\geq 6</math></b>	47 (19)	29 (62)
<b><math>\geq 8</math></b>	37 (15)	23 (62)
<b><math>\geq 9</math></b>	11 (4)	8 (72)
<b><math>\geq 11</math></b>	7 (3)	6 (86)

**Supplementary Table 3:** Complete overview of results from systematic review on investigated predictors for drug survival, drawn from 29 full-text articles [5-7, 38-41, 50-72] and 18 congress abstracts [8, 42-44, 73-86]. Predictors were drawn from multivariable models if possible and a p-value <0.05 was considered an effect. A predictor from a study was assigned to either one of three categories: associated with continuation, discontinuation or not associated. Per category, the number of studies that mention the particular predictor is shown. The association with survival (i.e. conclusion) is based on the consistency of the predictor in predicting the 'same direction' in more than 1/3rd of the studies.

Item (original articles) & (abstracts)	Original Articles Association with TNFi drug survival					Original Articles + Abstracts Association with TNFi drug survival				
	No. studied this item	Contin.	No assoc.	Dis- contin.	Concl.	No. studied this item	Contin.	No assoc.	Dis- contin.	Concl.
<b>Demographic and clinical parameters</b>										
Age* [7, 38, 40, 41, 50, 52, 55-57, 60-66, 68, 71, 72] & [8, 43, 73, 76, 83]	19	0	16	3	-	24	1	19	4	-
Female gender [38, 40, 41, 50, 52, 55-57, 59-66, 71, 72] & [76, 84]	18	1	16	1	-	20	1	17	2	-
Disease-duration [7, 38, 50, 55, 56, 59, 60, 62-65, 72] & [76, 83]	12	2	10	0	-	14	2	11	1	-
Comorbidity § [7, 52, 56, 57, 59, 60, 63, 64, 72] & [8, 84]	9	1	7	1	-	11	1	8	2	-
RF positivity/levels [7, 40, 50, 57, 64, 65, 71]	7	0	7	0	-					
Current smoking [5-7] & [8]	3	0	2	1	-	4	0	2	2	†
Pack years smoking [6, 7, 38]	3	0	1	2	†					
Ethnicity § [38, 64] & [79]	2	0	1	1	-	3	0	1	2	†
BMI § [55, 61]	2	0	1	1	-					
Nodules [7, 38]	2	0	2	0	-					
Erosions [7, 38]	2	0	2	0	-					
ACPA positivity [57, 65]	2	0	2	0	-					
High education level [50, 66] & [8]	2	1	1	0	-					
Weight [55] & [76]	1	0	1	0	-	2	0	2	0	-
Length [55]	1	0	1	0	-					
Extra articular manifestations [50]	1	0	1	0	-					
Employed [64]	1	0	1	0	-					
Socioeconomical status [50]	1	0	1	0	-					
Full pharmacy coverage [50]	1	1	0	0	-					

<b>Assessments</b>										
<b>DAS28*</b> [7, 38-40, 54, 55, 57, 59-63, 65, 67, 72] & [82]	15	1	11	3	-	16	1	12	3	-
<b>HAQ</b> (1;8;10;14-18;20-22) & [76, 84]	11	0	8	3	-	13	0	8	5	†
<b>SJC</b> [7, 38, 50, 55, 63, 65] & [42, 76]	6	0	6	0	-	8	0	7	1	-
<b>ESR</b> [7, 50, 57, 59, 60, 63] & (37)	6	1	4	1	-	7	1	5	1	-
<b>TJC</b> [7, 38, 50, 55, 63, 65]	6	0	4	2	-					
<b>CRP</b> [7, 38, 55, 57, 62, 65] & [76]	6	1	5	0	-	7	1	6	0	-
<b>ANA positivity</b> [65, 71] & [85]	2	0	2	0	-	3	0	3	0	-
<b>RADAI</b> [38, 40] & [84]	2	0	1	1	-	3	0	1	2	-
<b>VAS-general health</b> [7, 65]	2	0	2	0	-					
<b>Patient global assessment</b> [38, 55]	2	0	2	0	-					
<b>VAS-pain</b> [7] & [8]	1	0	1	0	-	2	0	2	0	-
<b>Physician global assessment</b> [38] & [76]	1	0	0	1	-	2	0	1	1	-
<b>HADS distress score</b> [7]	1	0	0	1	-					
<b>Fatigue</b> [38]	1	0	1	0	-					
<b>Physical activity</b> [38]	1	1	0	0	-					
<b>CDAI</b> [64]	1	0	1	0	-					
<b>FANA positivity</b> [53]	1	0	1	0	-					
<b>anti-ssDNA positivity</b> [53]	1	0	0	1	-					
<b>anti-dsDNA positivity</b> [65]	1	0	1	0	-					
<b>IgG</b> [65]	1	0	1	0	-					
<b>anti-Ro/SSA positivity</b> [65]	1	0	0	1	-					
<b>anti-La/SSB positivity</b> [65]	1	0	1	0	-					
<b>HTLV-I-positivity</b> [70]	1	0	0	1	-					
<b>CF-USD in synovium</b> [55]	1	1	0	0	-					
<b>Peridontitis &lt;5 year before start</b> [51]	1	0	0	1	-					
<b>Physical function</b> [8]	0	0	0	0	-	1	0	1	0	-
<b>Morning stiffness</b> [76]	0	0	0	0	-	1	0	1	0	-
<b>Treatments</b>										
<b>Concomitant MTX</b> [38, 52, 54-59, 62, 63, 67, 69, 71, 72] & [8, 42, 75-77, 86]	14	6	6	2	‡	20	10	7	3	‡
<b>Concomitant GCs</b> [38, 40, 54, 55, 57, 59, 60, 63, 64, 72] & [42, 76, 80, 81, 83, 84]	10	1	8	1	-	16	1	8	7	†

<b>Number of previously used DMARDs*</b> [7, 50, 56, 57, 60, 62, 63, 65]	8	0	6	2	-					
<b>Any concomitant csDMARD</b> [39, 58-60, 66, 69] & [74, 83, 86]	6	5	0	1	‡	9	8	0	1	‡
<b>Number previously used bDMARDs*§</b> [38-41, 50] & [42-44]	5	0	1	4	†	8	0	1	7	†
<b>Concomitant SSZ</b> [55, 69, 72] & [8]	3	0	2	1	-	4	1	2	0	-
<b>Concomitant NSAID</b> [38, 56] & [76]	2	1	1	0	-	3	1	1	1	-
<b>Current MTX dose (mg/week)</b> [65, 72]	2	0	2	0	-					
<b>Current GC dose (mg/week)</b> [57, 65]	2	0	2	0	-					
<b>Concomitant LEF</b> [69, 72]	2	0	1	1	-					
<b>Concomitant ciclosporin A</b> [71, 72]	2	0	2	0	-					
<b>Concomitant HCQ</b> [72]	1	1	0	0	-	2	2	0	0	-
<b>Any prior csDMARD</b> [38]	1	1	0	0	-					
<b>Cumulative MTX use, yrs</b> [38]	1	1	0	0	-					
<b>Concomitant AZA</b> [72]	1	0	1	0	-					
<b>Concomitant D-PEN</b> [72]	1	0	1	0	-					
<b>Previous use of LEF</b> [76]	0	0	0	0	-	1	0	0	1	-
<b>Season of treatment initiation</b> [78]	0	0	0	0	-	1	0	1	0	-

Legend:

\* combination of numerical and categorical values for this parameter

§ different categories of parameters per study investigated

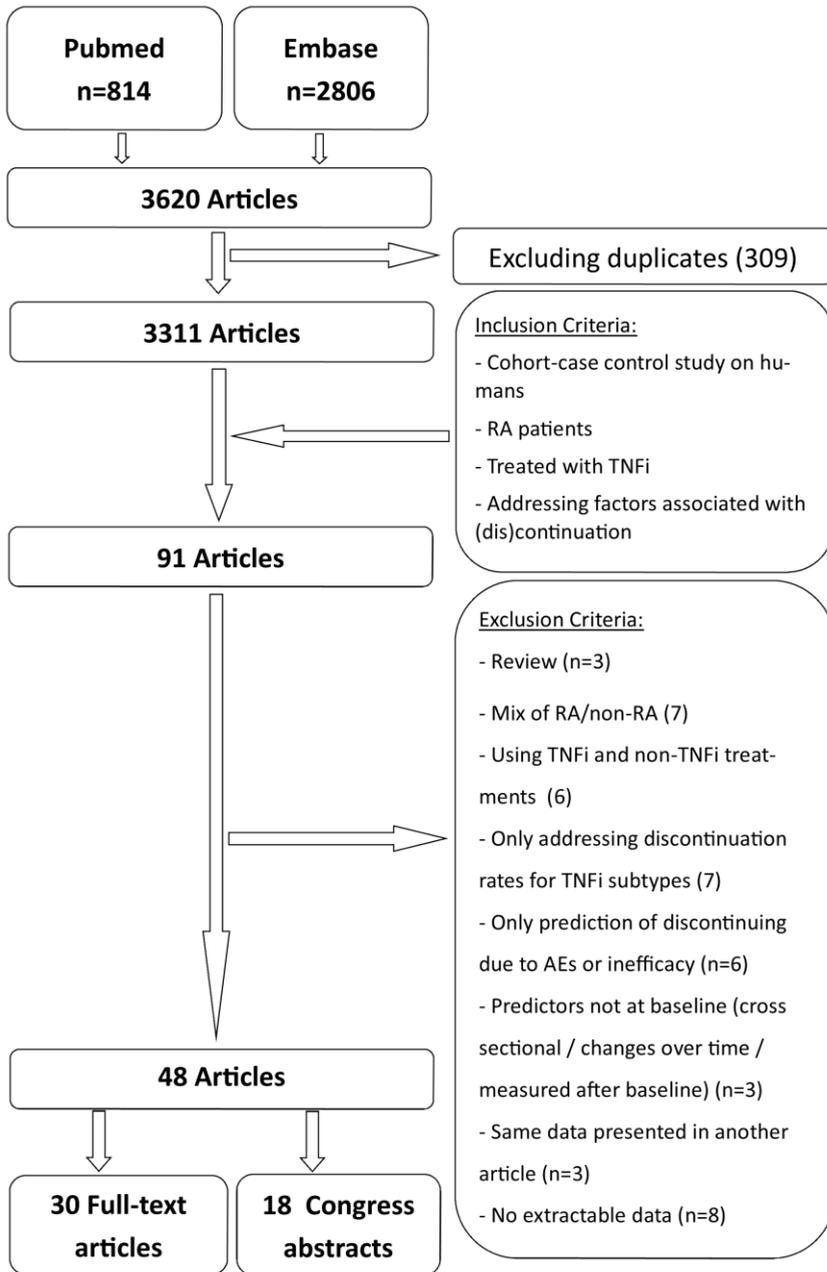
‡ predicting continuation (associated in >1/3rd of studies with continuation)

- no association with TNFi drug survival

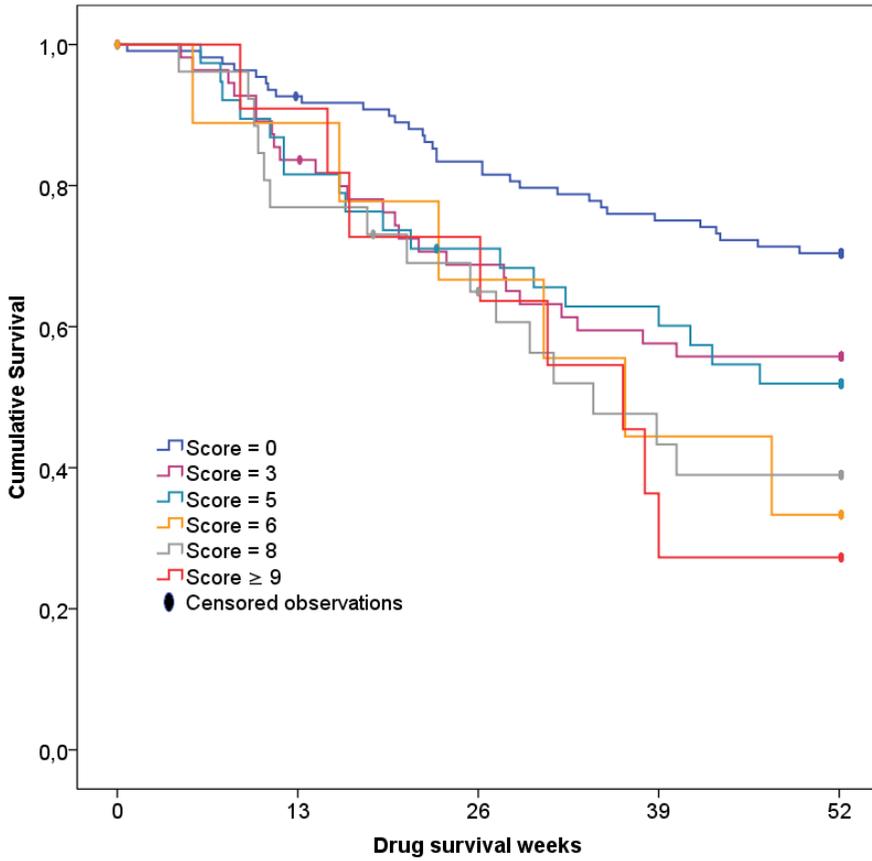
† predicting discontinuation (associated in >1/3rd of studies with discontinuation)

Abbreviations: ANA, anti-nuclear antibodies; anti-ss/dsDNA, anti single stranded/double stranded DNA antibodies; AZA, azathioprine; bDMARD, biological DMARD; BMI, body mass index; CDAI, clinical disease activity index; CF-USD, color fraction measured by ultrasound Doppler; CRP, C-reactive protein; csDMARD, conventional synthetic DMARD; D-PEN, D-penicillamine; DAS28, disease activity score based on 28 joints; DMARDs, disease modifying antirheumatic drugs; ESR, erythrocyte sedimentation rate; FANA, fluorescent antinuclear antibodies; GCs, glucocorticoids; HADS, hospital anxiety and depression scale; HAQ, health assessment questionnaire; HCQ, hydroxychloroquine; HLTV-I, human T lymphotropic virus type I; IgG, immunoglobulin G; LEF, leflunomide; MTX, methotrexate; NSAID, nonsteroidal anti-inflammatory drugs; RADAI, rheumatoid arthritis disease activity index; RF, rheumatoid factor; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count; VAS, visual analogue scale.

**Supplementary Figure 1:** Flow chart of systematic review and selection of original research articles for data extraction.



**Supplementary Figure 2:** TNFi survival curves for the selected cut-off of  $\geq 9$  (red) and lower scores in complete cohort (n=252).



## REFERENCES

### (supplementary only)

1. Agarwal SK, Glass RJ, Shadick NA, Coblyn JS, Anderson RJ, Maher NE et al. Predictors of discontinuation of tumor necrosis factor inhibitors in patients with rheumatoid arthritis. *J Rheumatol* 2008;35 9: 1737-44.
2. Bertoli AM, Strusberg I, Baravalle M, Betelu Z, Calas D, Morales L et al. Rate and causes of infliximab discontinuation in patients with rheumatoid arthritis in a private clinical practice. *J Clin Rheumatol* 2008;14 6: 313-7.
3. Chatzidionysiou K, Kristensen LE, Eriksson J, Askling J, van Vollenhoven R, Group FTA. Effectiveness and survival-on-drug of certolizumab pegol in rheumatoid arthritis in clinical practice: results from the national Swedish register. *Scand J Rheumatol* 2015;18: 1-7.
4. Chen HH, Chen DY, Lai KL, Chen YM, Chou YJ, Chou P et al. Periodontitis and etanercept discontinuation risk in anti-tumor necrosis factor-naïve rheumatoid arthritis patients: a nationwide population-based cohort study. *J Clin Rheumatol* 2013;19 8: 432-8.
5. Cho SK, Sung YK, Choi CB, Bae SC. Impact of comorbidities on TNF inhibitor persistence in rheumatoid arthritis patients: an analysis of Korean National Health Insurance claims data. *Rheumatol Int* 2012;32 12: 3851-6.
6. Comby E, Tanaff P, Mariotte D, Costentin-Pignol V, Marcelli C, Ballet JJ. Evolution of antinuclear antibodies and clinical patterns in patients with active rheumatoid arthritis with long-term infliximab therapy. *J Rheumatol* 2006;33 1: 24-30.
7. Du Pan SM, Dehler S, Ciurea A, Ziswiler HR, Gabay C, Finckh A. Comparison of drug retention rates and causes of drug discontinuation between anti-tumor necrosis factor agents in rheumatoid arthritis. *Arthritis Rheum* 2009;61 5: 560-8.
8. Ellegaard K, Christensen R, Torp-Pedersen S, Terslev L, Holm CC, Konig MJ et al. Ultrasound Doppler measurements predict success of treatment with anti-TNF- $\alpha$ ; drug in patients with rheumatoid arthritis: a prospective cohort study. *Rheumatology (Oxford)* 2011;50 3: 506-12.
9. Figueiredo IT, Morel J, Sany J, Combe B. Maintenance and tolerability of infliximab in a cohort of 152 patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2008;26 1: 18-23.
10. Finckh A, Simard JF, Gabay C, Guerne PA. Evidence for differential acquired drug resistance to anti-tumour necrosis factor agents in rheumatoid arthritis. *Ann Rheum Dis* 2006;65 6: 746-52.
11. Fisher A, Bassett K, Wright JM, Brookhart MA, Freeman H, Dormuth CR. Comparative persistence of the TNF antagonists in rheumatoid arthritis - A population-based cohort study. *PLoS One* 2014;9 8: e105193.
12. Frazier-Mironer A, Dougados M, Mariette X, Cantagrel A, Deschamps V, Flipo RM et al. Retention rates of adalimumab, etanercept and infliximab as first and second-line biotherapy in patients with rheumatoid arthritis in daily practice. *Joint Bone Spine* 2014;81 4: 352-9.
13. Gibofsky A, Cannon GW, Harrison DJ, Joseph GJ, Bitman B, Chaudhari S et al. Discontinuation of disease-modifying anti-rheumatic drugs and clinical outcomes in the Rheumatoid Arthritis DMARD Intervention and Utilisation Study 2 (RADIUS 2). *Clin Exp Rheumatol* 2015;33 3: 297-301.
14. Iannone F, Gremese E, Atzeni F, Biasi D, Botsios C, Cipriani P 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: 1179-84.
15. Iannone F, Gremese E, Gallo G, Sarzi-Puttini P, Botsios C, Trotta F et al. High rate of disease remission in moderate rheumatoid arthritis on etanercept therapy: Data from GISEA, the Italian biologics register. *Clin Rheumatol* 2014;33 1: 31-7.
16. Iannone F, Fanizzi R, Notarnicola A, Scioscia C, Anelli MG, Lapadula G. Obesity reduces the drug survival of second line biological drugs following a first TNF-(alpha) inhibitor in rheumatoid arthritis patients. *Joint Bone Spine* 2015;82 3: 187-91.
17. Kristensen LE, Saxne T, Nilsson JA, Geborek P. Impact of concomitant DMARD therapy on adherence to treatment with etanercept and infliximab in rheumatoid arthritis. Results from a six-year observational study in southern Sweden. *Arthritis Res Ther* 2006;8 6: R174.
18. Marchesoni A, Zaccara E, Gorla R, Bazzani C, Sarzi-Puttini P, Atzeni F et al. TNF-(alpha) antagonist survival rate in a cohort of rheumatoid arthritis patients observed under conditions of standard clinical practice. *Ann N Y Acad Sci* 2009;1173: 837-46.
19. Markenson JA, Gibofsky A, Palmer WR, Keystone EC, Schiff MH, Feng J et al. Persistence with anti-tumor necrosis factor therapies in patients with rheumatoid arthritis: Observations from the RADIUS registry. *J Rheumatol* 2011;38 7: 1273-81.

20. Martinez-Santana V, Gonzalez-Sarmiento E, Calleja-Hernandez M, Sanchez-Sanchez T. Comparison of drug survival rates for tumor necrosis factor antagonists in rheumatoid arthritis. *Patient Prefer Adherence* 2013;7: 719-27.
21. Matsudaira R, Tamura N, Sekiya F, Ogasawara M, Yamanaka K, Takasaki Y. Anti-Ro/SSA antibodies are an independent factor associated with an insufficient response to tumor necrosis factor inhibitors in patients with rheumatoid arthritis. *J Rheumatol*. 2011;2346-54.
22. Matthey DL, Dawes PT, Hassell AB, Brownfield A, Packham JC. Effect of psychological distress on continuation of anti-tumor necrosis factor therapy in patients with rheumatoid arthritis. *J Rheumatol* 2010;37 10: 2021-4.
23. Neovius M, Arkema EV, Olsson H, Eriksson JK, Kristensen LE, Simard JF et al. Drug survival on TNF inhibitors in patients with rheumatoid arthritis: comparison of adalimumab, etanercept and infliximab. *Ann Rheum Dis* 2015;74 2: 354-60.
24. Ostergaard M, Unkerskov J, Linde L, Krogh NS, Ravn T, Ringsdal VS et al. Low remission rates but long drug survival in rheumatoid arthritis patients treated with infliximab or etanercept: results from the nationwide Danish DANBIO database. *Scand J Rheumatol* 2007;36 2: 151-4.
25. Radovits BJ, Kievit W, Fransen J, Van De Laar MA, Jansen TL, Van Riel PL et al. Influence of age on the outcome of antitumor necrosis factor alpha therapy in rheumatoid arthritis. *Ann Rheum Dis* 2009;68 9: 1470-3.
26. Soderlin MK, Petersson IF, Geborek P. The effect of smoking on response and drug survival in rheumatoid arthritis patients treated with their first anti-TNF drug. *Scand J Rheumatol* 2012;41 1: 1-9.
27. Soliman MM, Ashcroft DM, Watson KD, Lunt M, Symmons DP, Hyrich KL. 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: 583-9.
28. Umekita K, Hidaka T, Miyauchi S, Ueno S, Kubo K, Takajo I et al. Treatment with anti-tumor necrosis factor biologic agents in human T lymphotropic virus type I-positive patients with rheumatoid arthritis. *Arthritis Care Res* 2014;66 5: 788-92.
29. Voulgari PV, Alamanos Y, Nikas SN, Bougias DV, Temekonidis TI, Drosos AA. Infliximab therapy in established rheumatoid arthritis: an observational study. *Am J Med* 2005;118 5: 515-20.
30. Yang CT, Kuo CF, Luo SF, Yu KH. Discontinuation of anti-TNF-alpha therapy in a Chinese cohort of patients with rheumatoid arthritis. *Clin Rheumatol* 2012;31 11: 1549-57.
31. Cho SK, Sung YK, Kim D, Won S, Choi JI, Choi CB et al. Safety of TNF inhibitors and predictors of its discontinuation in elderly patients with rheumatoid arthritis. *Ann Rheum Dis* 2014;73 SUPPL. 2 June 2014: 625.
32. Choquette D, Sauvageau D, Haraoui B, Raynauld J-P. Better retention rate at 5 years of anti-tnf agents used in conjunction with methotrexate over time in patients with rheumatoid arthritis: Real-life data from rhumadata computerized database. *Arthritis Rheum* 2012;64 SUPPL. 10: S198-S9.
33. Choquette D, Thorne JC, Kellsall JT, Zimmer M, Starr M, Sheriff MK et al. Methotrexate use at infliximab initiation and impact on treatment outcomes: An analysis from a canadian registry. *Arthritis Rheum* 2013;65 SUPPL. 10: S211-S2.
34. Faraawi R, Hack F, Bensen W, Choquette D, Rahman P. Prior use of leflunomide increases the hazard ratio of discontinuation of infliximab in a Canadian open label prospective study with overall drop-out rates similar to other biologic therapies for the treatment of rheumatoid arthritis. *J Rheumatol* 2009;36 11: 2571-2.
35. Favalli EG, Biggoggero M, Pregolato F, Becciolini A, Penatti AE, Marchesoni A et al. The 12-years retention rate of the first-line TNF-inhibitor in the treatment of rheumatoid arthritis: Real-life data from a local registry. *Arthritis Rheumatol* 2014;66 SUPPL. 10: S1038.
36. Gibofsky A, Saunders KC, Ganguli A, Cifaldi M, Grant S, Clewell J et al. Persistence and predictors of persistency of adalimumab among rheumatoid arthritis (RA) patients in a us registry. *Arthritis Rheum* 2012;64 SUPPL. 10: S165-S6.
37. Javaid MK, Batra R, Chana J, Round G, Judge A, Taylor P et al. Latitude but not season of initiation predicts clinical response to tnf therapy in patients with rheumatoid arthritis: the bsr biologics register-RA. *Rheumatology (Oxford)* 2013;52 SUPPL. 1 4: 2013-i82.
38. Kumar K, Situnayake D, Nightingale P, Deeming A, Buckley C, De PP et al. Differences in the efficacy and tolerability of anti-TNF therapies in patients with RA from diverse ethnic groups. *Rheumatology (Oxford)* 2014;53 SUPPL. 1: i91.

39. Lee SS, Lee KE, Park DJ. Drug survival rates of anti-tumor necrosis factor therapies in patients with rheumatoid arthritis. *Ann Rheum Dis* 2013;72 SUPPL. 3 June 2013: A875.
40. Lee J, Park D, Lee S. Drug survival of anti-tumor necrosis factor therapies in RA and AS. *Int J Rheum Dis* 2014;17 SUPPL. 1: 104.
41. Lie E, Lillegraven S, Fagerli KM, Uhlig T, Kvien TK. TNF inhibitor treatment in rheumatoid arthritis (RA) patients with moderate versus high disease activity at baseline: A comparison of utility gains, response and remission rates. *Arthritis Rheum* 2012;64 SUPPL. 10: S220.
42. McWilliams DF, Walsh DA. Factors predicting pain and early discontinuation of tumour necrosis factor- $\alpha$ -inhibitors in people with rheumatoid arthritis: Results from the British society for rheumatology biologics register. *Ann Rheum Dis* 2014;73 SUPPL. 2: 602.
43. Pocock JM, Albreshni S, Slack R, Polydoropoulou V, O'Reilly D, Ostor AJ. A comparison of drug survival times for infliximab, adalimumab and etanercept. *Rheumatology (Oxford)* 2009;48 SUPPL. 1: i69.
44. Raffeiner B, Botsios C, Sfriso P, Carletto A, Biasi D, Gremese E et al. Predictors, features and effects of first biologic switch in rheumatoid arthritis within gisea register: Italian 10-year experience. *Arthritis Rheum* 2011;63 SUPPL. 1 S10: 1201.
45. Ramiro S, Wolfe F, Michaud KD. Predictors of biologic response modifiers discontinuation in 2, 486 RA patients. *Arthritis Rheum* 2009;60 SUPPL. 10: 1020.
46. Saevarsdottir S, Santacatterina M, Turesson C, Forsblad H, Jacobsson L, Lindblad S. Clinical characteristics and outcome of golimumab treatment differs between bio-naive and patients previously exposed to biologicals. Nationwide results on Rheumatoid Arthritis (RA), Psoriatic Arthritis (PSA), ankylosing spondylitis (AS) and Other Spondylarthritides (SPA). *Arthritis Rheum* 2014;66 SUPPL. 10: S705.
47. Sung YK, Cho SK, Kim D, Won S, Choi JI, Choi CB et al. Antinuclear antibody positivity in RA patients and its impact on persistency of TNF inhibitors. *Ann Rheum Dis* 2014;73 SUPPL. 2 June 2014: 629.
48. Terabe K, Kojima T, Kaneko A, Hirano Y, Fujibayashi T, Hattori Y et al. The effectiveness of adalimumab concomitant with disease-modifying antirheumatic drugs other than methotrexate in rheumatoid arthritis. *Ann Rheum Dis* 2013;71 SUPPL. 3 June 2013: S1013-14.



## Chapter 4

### **Can baseline serum microRNAs predict response to TNF-alpha inhibitors in rheumatoid arthritis?**

Bart V.J. Cuppen<sup>1</sup>, Marzia Rossato, Ruth D.E. Fritsch-Stork, Arno N. Concepcion, Yolande Schenk, Johannes W.J. Bijlsma, Timothy R.D.J. Radstake, Floris P.J.G. Lafeber, on behalf of all SRU investigators

## **ABSTRACT**

### **Objective**

In rheumatoid arthritis, prediction of response to TNF-alpha inhibitor (TNFi) treatment would be of clinical value. This study aims to discover miRNAs that predict response and aims to replicate results of two previous studies addressing this topic.

### **Methods**

From the observational BiOCURA cohort, 40 adalimumab (ADA) and 40 etanercept (ETN) treated patients were selected to enter the discovery cohort and baseline serum profiling on 758 miRNAs was performed. The added value of univariately selected miRNAs ( $p < 0.05$ ) over clinical parameters in prediction of response was determined by means of the area under the receiving operator curve (AUC-ROC). Validation was performed by TaqMan single qPCR assays in 40 new patients.

### **Results**

Expression of miR-99a and miR-143 predicted response to ADA, and miR-23a and miR-197 predicted response to ETN. The addition of miRNAs increased the AUC-ROC of a model containing only clinical parameters for ADA (0.75 to 0.97) and ETN (0.68 to 0.78). In validation, none of the selected miRNAs significantly predicted response. miR-23a was the only overlapping miRNA compared to the two previous studies, however inversely related with response in one of these studies. The reasons for the inability to replicate previously proposed miRNAs predicting response to TNFi and replicate those from the discovery cohort were investigated and discussed.

### **Conclusion**

To date, no miRNA consistently predicting response to TNFi therapy in RA has been identified. Future studies on this topic should meet a minimum of standards in design that are addressed in this study, in order to increase the reproducibility.

## INTRODUCTION

Rheumatoid arthritis (RA) is a chronic, disabling disease that mainly affects the synovial joints, with a prevalence of 0.5-1.0% in the Western countries [1, 2]. The introduction of TNF- $\alpha$ -inhibiting therapy (TNFi), such as adalimumab (ADA) and etanercept (ETN), has dramatically improved the outlook for RA patients. Nevertheless, a substantial proportion of patients (approximately 30-40%) fail to respond to TNFi therapy [3, 4]. As we cannot predict before initiation of therapy which patients will be non-responders [5], TNFi treatment is administered in a trial and error approach. However in the time frame from initiation of therapy until response can be judged, which is usually three to six months later, non-responding patients suffer from uncontrolled disease with possible joint damage and the potential harmful side effects from treatment. The challenge is therefore to identify responders and non-responders to TNFi beforehand, so that TNFi use or considering alternatives can be encouraged.

microRNAs (miRNAs) are a large family of highly conserved noncoding genes that play a fundamental role in biological processes by controlling protein expression [6-8]. miRNAs execute these actions by binding to protein-coding messenger RNAs (mRNAs), resulting in translational repression or mRNA degradation [7]. Besides intracellularly, miRNAs are also found in several biological fluids, including saliva, plasma, serum and urine, either circulating in conjunction with specific carrier proteins or enclosed in extracellular vesicles [9, 10]. Exploring the use of circulating miRNAs as biomarkers for diseases has gained momentum in recent years because of the easy accessibility, the associations with specific disease conditions and their good stability [10, 11]. In RA, a systemic inflammatory disease primarily manifesting in the joints, biomarkers in the circulation would intuitively not be the most relevant compartment. However, the levels of miRNAs are frequently higher in the circulation than in the synovial fluid and correlate with disease activity in RA, indicating that the systemic compartment provides a useful compartment for studying the ongoing pathophysiological processes [12]. In addition, abnormal expression of both synovial and systemic miRNAs have been linked to disease activity and pathogenesis, even though their direct targets are not always known [13-19]. Three recent studies focused on the prediction of response to therapy in RA by using circulating miRNAs, of which two

investigated response to therapy with TNFi [20, 21] and one to rituximab [22]. According to these studies, promising predictors for TNFi therapy were miR-22 [20], miR-23a [21], miR-223 [21] and miR-886 [20]. Circulating miR-23a seems of particular interest, since it was the only identified candidate biomarker that was overlapping among both studies in univariate analyses. However, upregulation of miR23a was found in whole blood[20], whereas a downregulation was found in serum [21] of future responders.

In this study we explored the serum miRNAs associated with good and bad response to TNFi therapy, in order to replicate the results that have been published before. In addition to the previous studies performed, we involved clinical parameters in the prediction and attempted to validate the miRNAs and prediction models in a separate cohort.

## **METHODS**

### **Clinical data collection**

Patients initiating ADA or ETN therapy were selected from the “Biologicals and Outcome Compared and predicted Utrecht region in Rheumatoid Arthritis” (BiOCURA) study. BiOCURA is an observational cohort, in which RA patients eligible for biological treatment according to regular clinical practice were enrolled and followed after start of treatment, in one academic hospital and seven regional hospitals in the Netherlands (see Acknowledgements). Re-inclusion after switching to a different biological treatment was possible, at which patients re-entered baseline again. The study was approved by the local ethics committee of the University Medical Center Utrecht and the institutional review boards of the participating centers, and was performed in accordance with the Declaration of Helsinki. Informed consent was obtained from each patient.

Trained nurses gathered all data during a dedicated visit, which included all clinical parameters, joint counts and collection of blood. Visits were scheduled at baseline (before initiation) and after three, six and twelve months of treatment. Disease activity was assessed using the disease activity score based on a 28-joint count (DAS28) [23] and subsequently the European League Against Rheumatism (EULAR) response was calculated [24]. This study design, allowed the determination of a clinical response of each patient, on the basis of three EULAR responses over the course of one year.

## Patient selection

Two separate cohorts were composed from the available patients in the BiOCURA study: a discovery cohort was used to screen the expression of a full panel of 758 miRNAs, while a validation cohort was used to test if the results found in the discovery phase were reproducible. The discovery cohort was formed by selecting the most extreme patients regarding clinical response, among all ADA and ETN treated patients included between June 2009 (start of BiOCURA) and October 2012 (n=74 ADA and n=68 ETN). The top responding patients (n=20 for both ADA and ETN, from now on called “responders”), were identified by the selection of patients with the best three EULAR responses over the course of one year. The selection of bad responders (n=20 for both ADA and ETN, from now on called “non-responders”), was based on the most negative EULAR responses over the course of one year and/or (early) discontinuation of TNFi treatment due to inefficacy. Patients with a baseline DAS28 <2.6 were excluded from the analysis in order to reduce the chance that limited improvement in DAS28 resulted in misclassification as EULAR non-responders. For validation of results, responders (n=10 for both ADA and ETN) and non-responders (n=10 for both ADA and ETN) were selected using the same criteria as in the discovery cohort, among patients included from October 2012 until June 2015 (n=25 ADA and n=40 ETN). Since the validation cohort was smaller, relatively more patients were selected and the differences in clinical outcome between responders and non-responders were less extreme. The baseline characteristics for responders and non-responders are shown in Table 1 and for responders and non-responders split per cohort in Supplementary table 1. Additionally, the baseline characteristics for the discovery and validation cohorts are shown in Supplementary Table 2.

**Table 1: Baseline characteristics of responders and non-responders, split for treatment received.** RA patients were selected from the observational BiOCURA cohort based on treatment outcome over the course of one year after the start of treatment with either ADA or ETN. The presented clinical characteristics for responders and non-responders refer to the values present before treatment initiation. P-values of comparisons between responders and non-responders were calculated by means of an independent sample t-test, Mann-Whitney U test, fisher exact test (2\*2) or chi-square (>2\*2) based on the distribution of the clinical parameter. Bold p-values indicate significant associations (p<0.05).

Item	ADA (n=60)		p-value	ETN (n=60)		p-value
	Non-resp (n=30)	Resp (n=30)		Non-resp (n=30)	Resp (n=30)	
Female gender, n (%)	21 (70)	21 (70)	1.00	25 (83)	21 (70)	0.36
Age, mean years $\pm$ sd	54.4 $\pm$ 10.9	53.5 $\pm$ 12.7	0.76	58.3 $\pm$ 9.2	55.1 $\pm$ 10.5	0.22
Current smoker, n (%)	16 (53)	8 (27)	0.06	8 (27)	7 (23)	1.00
RF positivity, n (%)	16 (53)	21 (70)	0.29	20 (67)	22 (73)	0.78
ACPA positivity, n (%)	19 (63)	19 (63)	1.00	19 (63)	26 (87)	0.07
CRP, mg/l median (IQR)	5.2 (1.6-10.5)	5.5 (2.0-12.3)	0.78	4.0 (2.0-9.0)	8.5 (4.0-18.3)	<b>0.03</b>
No. of previously used bDMARDs			0.48			1.00
0, n (%)	20 (67)	23 (78)		22 (73)	22 (73)	
1, n (%)	9 (30)	7 (23)		7 (23)	7 (23)	
2, n (%)	1 (3)	0 (0)		1 (3)	1 (3)	
Concomitant treatment, n (%)	29 (97)	29 (97)	1.00	27 (90)	29 (97)	0.61
MTX, n (%)	21 (70)	27 (90)	0.10	18 (60)	25 (83)	0.08
SSZ, n (%)	2 (7)	4 (13)	0.67	4 (13)	2 (7)	0.67
HCQ, n (%)	8 (27)	7 (23)	1.00	10 (33)	11 (37)	1.00
GC, n (%)	15 (50)	4 (13)	<b>0.01</b>	11 (37)	6 (20)	0.25
Baseline DAS28, mean $\pm$ sd	3.9 $\pm$ 1.4	4.7 $\pm$ 0.9	<b>0.01</b>	4.3 $\pm$ 1.2	4.6 $\pm$ 0.9	0.21
TJC, median (IQR)	5.0 (1.0-13.0)	7.0 (4.0-14.3)	0.35	6.5 (2.8-11.3)	5.0 (2.8-11.3)	0.87
SJC, median (IQR)	0.0 (0.0-4.0)	2.0 (0.0-4.0)	<b>0.03</b>	1.0 (0.0-3.3)	2.0 (0.8-4.0)	0.20
VAS-GH, mean $\pm$ sd	55.2 $\pm$ 23.8	63.8 $\pm$ 22.0	0.15	55.5 $\pm$ 22.8	55.1 $\pm$ 10.5	0.76
ESR, median mm/hr (IQR)	11.0 (3.8-26.0)	16.5 (9.0-32.0)	0.14	13.0 (5.8-33.8)	21.0 (14.3-39.5)	0.07

ACPA, anti-citrullinated protein antibody; ADA, adalimumab; bDMARDs, biological disease modifying antirheumatic drugs; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; ETN, etanercept; GC, glucocorticoid; HCQ, hydroxychloroquine; IQR, interquartile range; MTX, methotrexate; RF, rheumatoid factor; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count; VAS-GH, visual analogue scale general health.

## **Blood processing and RNA extraction**

Blood was collected in Vacutainer® SSTII-tubes (BD, Franklin Lakes, NJ, USA) and processed immediately after clotting. Samples were centrifuged for 10 min at 1500g at room temperature and serum was aliquoted and stored at -80°C until use. RNA was extracted from 240µl of serum using the miRcury RNA Isolation kit for Biofluids (Exiqon, Woburn, MA, USA), according to the manufacturer's instructions. During extraction, 300pg of a synthetic miRNA (*Arabidopsis thaliana* ath-miR-159a) was added to each sample as a spike-in to monitor technical variability along the isolation procedure and for later normalization.

## **miRNA profiling**

miRNA profiling was performed by TaqMan RT-qPCR on the OpenArray platform (Life Technologies, Carlsbad, CA, USA). This method allows the simultaneous analysis of 758 miRNAs, split into two equal pools (A and B). Manufacturer's instructions were followed with minor adjustments. Briefly, 2.5µl of isolated serum RNA was reverse-transcribed by using the miRNA multiplex RT primers pools, either v2.1 for pool A or v3.0 for pool B, and the TaqMan miRNA reverse transcription kit (Life Technologies). RT products were pre-amplified using the Megaplex PreAmp Primers pools A and B in the presence of the TaqMan PreAmp Master mix (Life Technologies), by using the following thermal cycler conditions: 10 min,95°C; 2 min,55°C; 2 min,72°C and 16 cycles of 15 sec,95°C and 4 min,60°C and one single cycle of 10 min,96°C. The miRNA OpenArray profiling was performed on the amplified cDNA, diluted to 1:40, with 0.1× TE buffer pH 8.0 and subsequently 1:2 by using the Taqman OpenArray Master Mix on the QuantStudio 12k flex Real-Time PCR system (Life Technologies).

miRNA profiling data was analyzed using the Relative Quantification application in the online accessible Thermo Fisher Cloud (<https://apps.thermofisher.com/apps/dashboard/>), using the relative threshold cycle (Crt) and the comparative threshold cycle method [25]. Briefly, miRNA expression was calculated after normalization by exogenous ath-miR-159a spike-in ( $\Delta\text{Crt} = \text{Crt mean target} - \text{Crt mean miR-159a}$ ). The relative fold change (FC) of each sample was determined by setting the FC of a random ADA or ETN non-responder sample at 1, and calculate the FC compared to this reference ( $\text{FC} = 2^{-\Delta\Delta\text{Crt}}$ , where  $-\Delta\Delta\text{Crt} =$

$\Delta\text{Crt}$  reference –  $\Delta\text{Crt}$  sample). Low expressed miRNAs, i.e. having  $\text{Crt}$  higher than 27 were set to 27, and samples with a low amplification quality (i.e. amplification score  $<1.24$ ) were excluded from the analysis.

### **Individual miRNA analysis**

miRNA-specific TaqMan Real-Time quantitative PCR (RT-qPCR) assays were purchased from Life Technologies for hsa-miR-23a-3p (ID 000399), hsa-miR-99a-5p (ID 000435), hsa-miR-143-3p (ID 002249), hsa-miR-197-3p (ID 000497) and for the exogenous control ath-miR-159a (ID 000338). From 2.5 $\mu\text{l}$  baseline serum RNA, cDNA was synthesized by using individual miRNA-specific RT primers contained in the miRNA assay in the presence of 3.3 U/ $\mu\text{l}$  MultiScribe RT enzyme (Life Technologies), by using the following thermal cycler conditions: 10 min, 4°C; 30 min, 16°C; 30 min, 42°C; and 5 min, 85°C. Circulating miRNA levels were quantified in duplicate from 3 $\mu\text{l}$  cDNA, with TaqMan Fast Advance Master Mix and specific primers of the miRNA assay, using the following amplification condition on the Quantstudio 12k flex Real-Time PCR system: 2 min, 50°C; 20 sec, 95°C; 40 cycles of 1 sec, 95°C; and 20 sec, 60°C. RT-qPCR data were calculated as described above, with the difference that baseline threshold cycles ( $\text{Ct}$ ) were used.

### **Statistical analyses**

Differential expression of miRNAs between responders and non-responders, was calculated separately for ADA and ETN by means of an independent sample t-tests on the  $-\Delta\Delta\text{Crt}/-\Delta\Delta\text{Ct}$ , with a threshold for significance of 0.05 (uncorrected p-value). The levels of differentially expressed miRNAs were plotted in GraphPad Prism (GraphPad, La Jolla, CA, USA) as FC of responders *versus* non-responders. Validation was considered successful when both the t-test was significant and plots of the FC showed the same direction (i.e. up/downregulation).

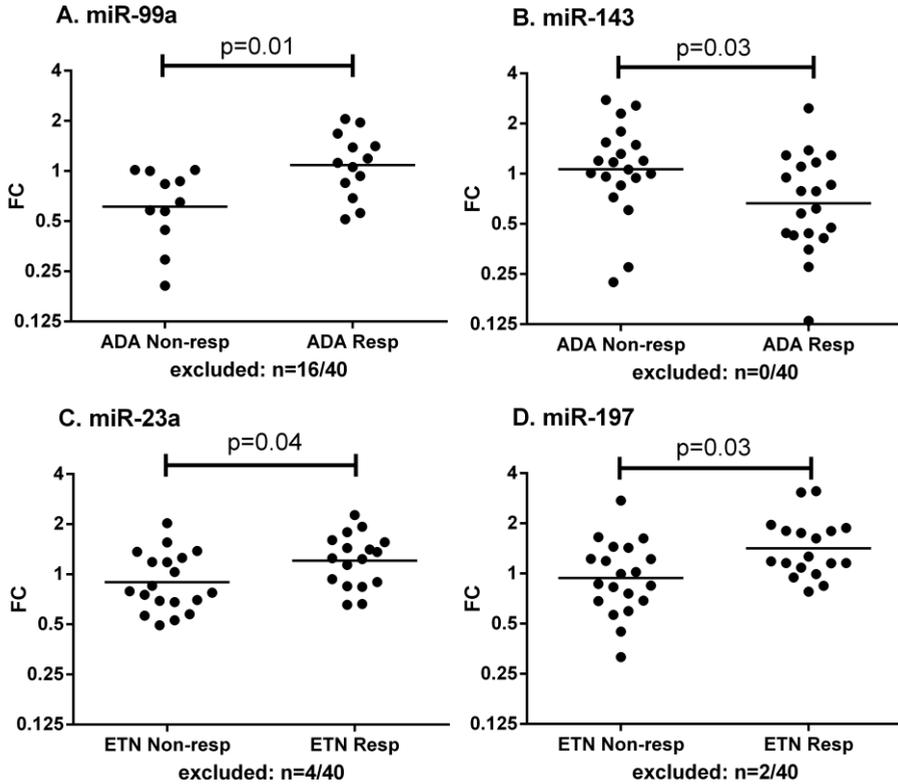
In order to determine the added value of the miRNAs over clinical parameters, we built two prediction models for each treatment, using multivariable logistic regression. The first model consisted of all baseline clinical parameters that were significantly different between responders and non-responders (the “clinical model”). The second model included the clinical parameters and the selected miRNAs ( $-\Delta\Delta\text{Crt}$  values) (the “combined

model"). Per model, the area under the receiver operating characteristic curve (AUC-ROC) was calculated as an indicator of the predictive ability. We considered an AUC-ROC of <0.7 limitedly, 0.7-0.8 moderately and >0.8 highly predictive of response. The sensitivity and specificity were calculated for the best cutoff value per model, according to Youden's index [26]. Evaluation of the added value of miRNAs was based on the increase of predictive abilities when switching from the clinical to the combined model. In order to validate the findings from multivariable analysis, the prediction rules of the clinical and combined models were applied in the validation cohort, thereby freezing the regression coefficients of the individual parameters from the original model. Again, the AUC-ROC, the sensitivity and specificity were calculated to interpret the added value of miRNAs over clinical parameters alone.

## **RESULTS**

### **Identification of miRNAs as predictor of TNFi response**

We analyzed the profile of miRNAs present in the circulation of responders versus non-responders with a broad panel of 758 miRNAs. In the discovery cohort (n=80), four miRNAs were significantly differentially expressed between responders and non-responders: high and low baseline levels of respectively miR-99a and miR-143 predicted response to ADA, while patients with high levels of miR-23a and miR-197 more frequently responded to ETN (Figure 1). Expression values of patients in the discovery cohort were also plotted for miRNAs proposed by the previous studies as predictors for response (Supplementary Figure 1). miR-23a in ETN treated patients was the only miRNA overlapping between this study and previous ones published.



**Figure 1:** miRNAs significantly differentially expressed in the discovery cohort. A large panel of miRNAs was measured using the OpenArray platform in serum of 40 ADA and 40 ETN treated patients. miRNAs showing significant differences ( $p < 0.05$ ) between responders and non-responders were selected as potential predictors. Among all analyzed, four miRNAs were selected as potential predictors. Levels of miRNAs in each individual patient are shown as the fold changes (FCs) for ADA (A & B) and ETN (C & D). The geometric mean per group is shown and p-values between responders and non-responders were calculated on the  $-\Delta\Delta C_{rt}$  using an independent sample t-test. Several patients were excluded from the analysis because of low amplification quality (scores  $< 1.24$ ): miR-99a (n=16), miR-143 (n=0), miR-23a (n=4) and miR-197 (n=2).

Since the measurement of miRNAs can be costly when incorporated in clinical practice, we wanted to rule out the possibility that the miRNAs identified do not increase the magnitude of prediction that is already possible based on clinical parameters. We therefore compared the predictive abilities of models based on clinical parameters alone, and clinical parameters together with the miRNA expression levels. The clinical characteristics that were used, were those that presented a significant difference between

responders and non-responders at baseline, namely the DAS28 ( $p=0.01$ ), swollen joint count (SJC,  $p=0.03$ ) and amount of glucocorticoid (GC) users ( $p=0.01$ ) for patients treated with ADA, and C-reactive protein (CRP,  $p=0.03$ ) for those treated with ETN (Supplementary Table File 1). The predictive properties of these models without and with miRNAs are shown in Table 2. The clinical model for ADA showed a moderate predictive value (AUC-ROC 0.75), that was increased by the addition of miR-99a and miR-143 in the combined model (AUC-ROC 0.97). For ETN, the CRP alone was only limitedly able to predict response (AUC-ROC 0.68), however, the predictive value increased by the addition of miR-23a and miR-197 in the combined model (AUC-ROC 0.78).

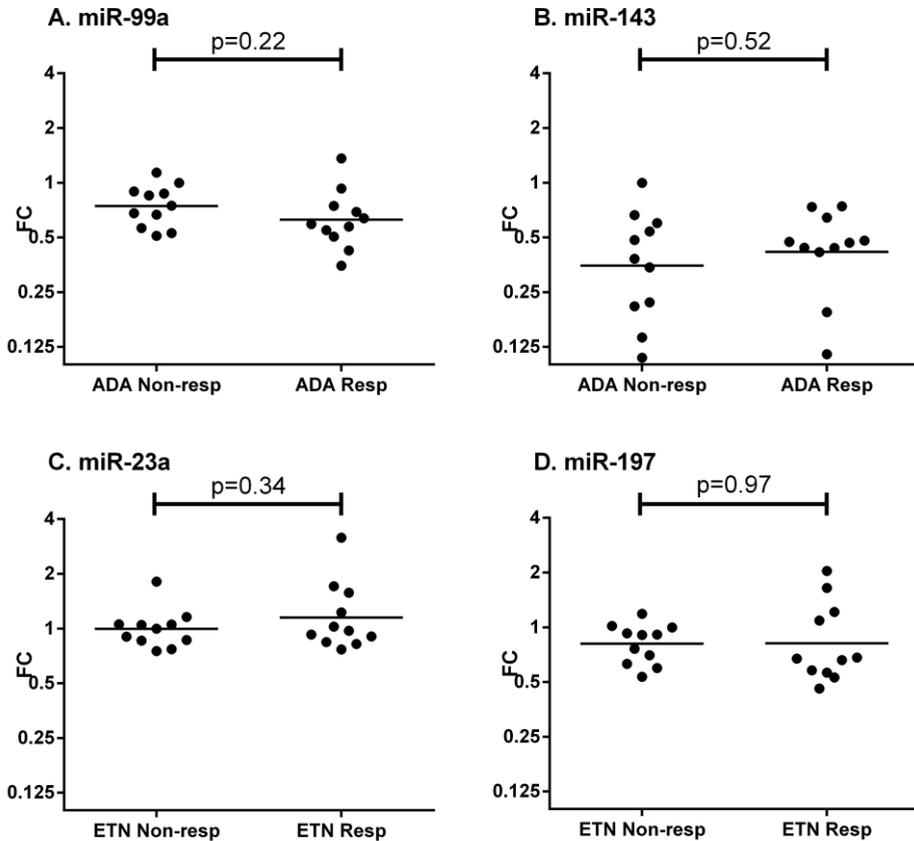
**Table 2:** Multivariable models for prediction of response to TNFi. Baseline clinical parameters of patients that were different between responders and non-responders ( $p<0.10$ ) were used to build a “clinical model”. In a “combined model”, the clinical parameters and miRNAs predictive for response were combined, in order to determine the additive value of miRNAs in the prediction of response. For ADA, a model containing the clinical parameters (the square root of) SJC, DAS28 and GC use was compared with a model containing these parameters and the level of circulating miRNAs associated with response to ADA, miR-99a and miR143 ( $-\Delta\Delta C_{rt}$  values). For ETN, the clinical model only contained the (logtransformed) CRP and the combined model also included miR-197 and miR23. Per model AUC-ROC is shown as an indicator of the predictive ability. A useless model would score 0.5, whereas a perfect model would score 1.0. The sensitivity (proportion of test positive among all responders) and specificity (proportion of all test negative among all non-responders) were shown for the best cut-off value per model, according to the Youden’s index.

TNFi	Model	Model content	AUC-ROC	Sens.	Spec.
ADA	Clinical	SJC, GC use, DAS28	0.75	80%	70%
	Clinical + miRNAs	SJC, GC use, DAS28, miR-99a, miR-143	0.97	92%	91%
ETN	Clinical	CRP	0.68	67%	75%
	Clinical + miRNAs	CRP, miR-197, miR-23a	0.78	80%	79%

ADA, adalimumab; AUC-ROC, area under the receiver operating characteristic curve; CRP, C-reactive protein; ETN, etanercept; GC, glucocorticoid; SJC, swollen joint count; TNFi, TNF- $\alpha$ -inhibitor.

Since replication in (prognostic) research is key to prove validity, we tried to confirm our results in an additional cohort of 40 patients. The differentially abundant miRNAs from the discovery cohort were analyzed in the validation cohort by using single RT-qPCR assays (Figure 2). None of the miRNAs could significantly predict the response to TNFi in the

validation cohort ( $p > 0.05$ ). For miR-99a and miR-143 in ADA users, inverse directions were seen compared to the results in the discovery cohort.



**Figure 2:** Validation of selected miRNAs. Using single assays, miRNAs selected in the discovery cohort were measured in an independent cohort of patients treated with ADA (A & B) (n=20) and ETN (C & D) (n=20). miRNAs were considered validated when showing the ‘same direction’ of variation as in the discovery cohort and a significant difference ( $p < 0.05$ ) between responders and non-responders. Shown are the fold changes (FCs) of the individual patients and the geometric means per group. P-values were calculated on the  $-\Delta\Delta C_{rt}$  using an independent sample t-test. No patients were excluded from the analysis because of low amplification scores.

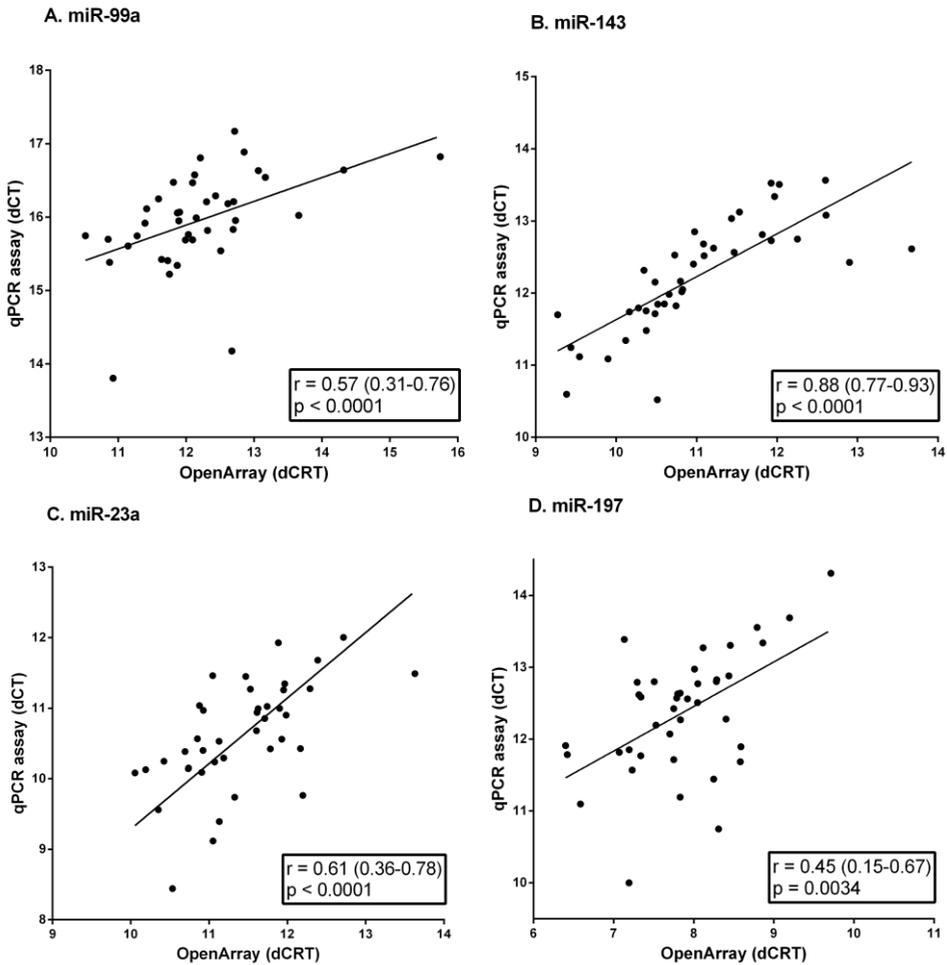
Multivariable analyses did not confirm the predictive abilities of the combined models found in the discovery phase (Supplementary Table 3). Contrary to what was observed in the discovery phase, application of the prediction model for ADA including only clinical parameters showed better predictive abilities than the corresponding combined model (AUC-ROC from 0.93 to 0.57). This is most likely explained by the inverse relationship of the miRNAs and response in the validation compared to the discovery cohort. For ETN, the addition of miRNA added to the prediction of response (AUC-ROC from 0.59 to 0.66), generating a model that was only limitedly predictive.

### **Factors that can contribute to the inability to validate findings in multiple cohorts**

It is of importance to investigate which factors are involved in the inability to validate findings in multiple cohorts, since these issues could also be applicable to other studies aiming to identify miRNAs predicting the response to TNFi therapy. Possible (combinations of) factors could be, but are not limited to, the usage of different miRNA detection methods, selection of false positive results in the discovery phase, or clinical parameters influencing the relationship between miRNA levels and response.

Despite that miRNA expression analysis in the discovery and validation phase were both based on the same method of detection, i.e. miRNA-specific retrotranscription combined with TaqMan-based RT-qPCR, the scale (high throughput *versus* single assay) of the techniques sufficiently varies. To evaluate whether these differences could impact the final result, we have performed a technical replication in all 40 ADA or 40 ETN samples from the discovery cohort using single assays for the four selected miRNAs (as described in the Methods section “Individual miRNA analysis”). Correlations of the results obtained by the profiling *versus* those measured by single assay were assessed by calculating the Spearman correlation ( $r$ ) between the normalized detection levels ( $\Delta C_{rt}$  and  $\Delta C_t$  respectively) without excluding samples based on amplification scores (Figure 3). The correlations of test-retest values ranged from 0.45 to 0.88 (all  $p < 0.0001$ ), which can be considered ranging from reasonable to good, thus demonstrating that the two analyses are concordant. Comparison of the single assay miRNA expression levels between responders *versus* non-responders confirmed that miR-143 was significantly lower in ADA-

responders, whereas the other miRNAs showed the same direction as in the profiling, though did not reach significance (Supplementary Figure 2).



**Figure 3:** Correlation between OpenArray and single-assay results. A technical replication of the four selected miRNAs was performed. Per miRNA, all ADA (A & B) or ETN (C & D) samples from the discovery cohort were re-analyzed using Taqman single miRNA assays. The normalized values for the OpenArray ( $\Delta$ Crt) and single assay ( $\Delta$ Ct) for all 40 samples was plotted and the Spearman correlation (r) was calculated.

To verify whether these results could be related to a false discovery rate, we recalculated the differential expression for all miRNAs in the discovery cohort while applying the Benjamini and Hochberg false discovery rate (B&H FDR), which showed corrected p-values of 1.00 for all miRNAs. Considering none of the miRNAs was significantly different after correction, there is the possibility that only false positive results were selected in the discovery phase.

Another possible explanation why we were unable to replicate the results from previous studies and our discovery cohort, is that clinical parameters interact with miRNA levels and these clinical parameters were not equally distributed between the cohorts. Differences in case-mix between cohorts that were unaccounted for (see Supplementary Table 2), would then lead to different estimations of each miRNA and response. Despite this, an adjustment for these clinical parameters would then give comparable estimations for the miRNAs involved. We investigated this theory by running a crude model of response including the specific miRNA only, and an adjusted model considering both the miRNA and clinical parameters, and run these models for the two cohorts analyzed (Table 3). Despite the adjustment, the odds ratios (OR) of these miRNAs for response were still (very) different between the discovery and validation cohort. This indicates that the clinical parameters do not explain why results could not be validated. On the other hand, these analyses showed that clinical parameters have a strong effect on the association between miRNA levels and the response to therapy, as indicated by the (relatively large) differences between crude ORs and adjusted ORs. Therefore, adjustment for clinical parameters will contribute to externalization of results to cohorts with a different case-mix, as is a common occurrence in a heterogeneous disease such as (established) RA. Considering that clinical parameters could affect the expression of miRNAs, miRNA levels may be, to a certain extent, a representation of patient's clinical characteristics. Therefore, we evaluated the correlation of clinical parameters and miRNAs levels, as measured by single assays, irrespective of response. The analysis revealed that all miRNAs associated with either CRP or erythrocyte sedimentation rate (ESR) (Supplementary Table 4). However, since the explained variance of each miRNA by clinical parameters was less than 35%, miRNA levels are not a complete reflection of clinical characteristics and can thus contain informative additional information.

**Table 3:** Influence of clinical parameters on the association of each miRNA to response. In order to test if clinical parameters influenced the association between miRNA levels and response, each univariately selected miRNA was first inserted in a logistic regression model on response (crude) in the discovery and validation cohort ( $-\Delta\Delta C_{rt}$  and  $-\Delta\Delta C_t$  values respectively). Then the baseline clinical parameters that were different between discovery and validation (Supplementary Table 2,  $p < 0.10$ ) were added to a separate model (adjusted). The parameters that were used to adjust the association of miR-99a and miR-143 with response to ADA were (logtransformed) CRP, DAS28, (the square root of) SJC and VAS-GH. The association of miR-23a and miR-197 with response to ETN was adjusted for age and (logtransformed) ESR. If clinical parameters would be the main cause for the inability to validate the findings, the adjusted ORs of the miRNA in discovery and validation should be comparable. The analyses showed that relationship between each miRNA and response was significantly influenced by clinical parameters in most cases, however, because the adjusted ORs between discovery and validation after correction are not comparable, the clinical parameters do not explain the found differences between the different cohorts.

miRNA	Cohort	Crude OR	Crude p-value	Adjusted OR	Adjusted p-value
<b>miR-99a</b>	Discovery	6.78	0.03	6.82	0.06
	Validation	0.32	0.28	0.85	0.91
<b>miR-143</b>	Discovery	0.45	0.04	0.39	0.04
	Validation	1.75	0.31	3.43	0.24
<b>miR-23a</b>	Discovery	4.08	0.03	3.82	0.05
	Validation	2.81	0.31	1.86	0.65
<b>miR-197</b>	Discovery	4.32	0.02	5.00	0.03
	Validation	1.41	0.68	0.99	0.99

ESR, erythrocyte sedimentation rate; OR, odds ratio; SJC, swollen joint count; VAS-GH, visual analogue scale general health.

Since the exact origin of circulating of miRNAs is unknown and blood cells have been proposed as a key source [27], we wanted to rule out the possibility that the serum levels of miRNAs are a reflection of the composition of circulating leukocytes. We therefore correlated the levels of serum miRNAs with the percentage of peripheral blood leukocyte subsets, as measured by flow cytometry in 20 randomly selected patients (Additional File, not included in this thesis). Out of all comparisons that were considering the surface markers CD3, CD4, CD8, CD14, CD19, CD45 and CD16 + 56 and the FC of all four miRNAs, one significant correlation was observed between the levels of miR-197, as measured in the profiling, and the percentage of natural killer-like T-cells ( $r=0.587$ ,  $p=0.008$ ). However, this association was not reproduced when considering the miRNA levels measured by the single assay ( $r=0.083$ ,  $p=0.831$ ). Even though we cannot exclude the contribution of other

rare cell subsets that were not identified, these results suggested that the levels of serum miRNAs in responders *versus* non-responders are not dependent on the composition of circulating leukocytes, thus making a different leukocyte composition an unlikely cause for the inconsistency between the two cohorts analyzed and the other studies.

## **DISCUSSION**

Prediction of TNFi response is needed for a more personalized approach in RA treatment. Since two previous studies have addressed this question and identified candidate miRNAs [20, 21], we aimed at verifying whether these could be validated in an independent cohort, and eventually, whether we could find new predictors. High values of circulating miR23a were univariately predictive of response in our study and one previous work [20], however, miR-23a was inversely related to response in a third study [21] and is thus not a consistent predictor. On the basis of the profiling results, four miRNAs showed an added value to clinical parameters in predicting response to TNFi. However, these miRNAs could not be validated in a separate cohort of consecutively included patients.

Several factors could have contributed to the inability to replicate findings from previous works and our own selection of miRNAs. A first possible contributing factor is technical variability of the techniques employed for the miRNA analysis. Within our study and the previous studies, pre-analytical and analytical protocols were standardized, as is considered to contribute to more reliable measurements in miRNA studies[28]. Correlations between the two protocols used to analyze miRNA expression were reasonably good, though could explain (some) difference in the outcome. Indeed, a significant differential expression between responders and non-responders could only be replicated for miR-143, which was the miRNA with the best amplification quality in the array among those selected (i.e. amplification score  $\geq 1.24$  in all samples). The previous studies used comparable techniques for biomarkers discovery: TaqMan single miRNA assays [21] and the same platform as in our discovery step (TaqMan OpenArray, Pool A)[20]. Furthermore, the OpenArray appears one of the most reliable high throughput technique for miRNA analysis [29], and it was successfully used for profiling of serum

miRNA in multiple studies by us [Chouri E. et al, manuscript in preparation] and others [30, 31]. Interestingly, miR-23a was positively related to response both in our discovery cohort and in a previous study that was also using the OpenArray as profiling platform [20]. On the contrary the absence of a relation or inverse relation with response was found in all cases employing TaqMan single assays [21]. Altogether, the analytical techniques and their technical variability are unlikely to be explanatory for all the differences within- and between the studies, though they might be a contributing factor.

The lack of correction for multiple testing in the discovery phase might have led to a subsequent (wrongful) selection of candidate miRNAs. However, a correction for multiple testing was not applied, because this could be too strict when trying to replicate findings already identified by others and would also have increased the chance of excluding potentially useful miRNAs (false negative results). In addition, a separate validation step will reveal which of the (less strictly) selected miRNAs has a true biological meaning, and thus compensates for the absence of correction for false discovery rate. Yet, the fact that the application of B&H FDR resulted in a p-value of 1.00 for all analyzed miRNAs is suggestive for the possibility that false positive miRNAs were selected. The identification of miRNAs in the previous studies might also be based on false positive results, since no correction for multiple testing was performed and results were not validated in a separate cohort. Not unimportant, if all proposed predictors so far are indeed false positives, this might entail an absence of biologically relevant miRNAs for the prediction of TNFi response in RA.

A third contributing factor for the inability to validate our findings could be related to differences in baseline clinical parameters, which could influence the relation between miRNAs and response. In additional analyses, however, we demonstrated that the clinical parameters were not able to explain why results could not be replicated between the discovery and the validation cohort. However, we did observe that the predictive values of miRNAs were dependent on clinical parameters (large difference in crude OR and adjusted OR) and that, to a certain extent, circulating miRNA levels are a reflection of clinical parameters. Since the other studies did not correct for clinical parameters, heterogeneity

in baseline characteristics of included patients might explain why miRNAs were not reproduced across studies. This is especially likely to have occurred if the heterogeneity across studies involves clinical characteristics relating to response, such as baseline DAS28, SJC, GC use and CRP, which were indeed the parameters that differed across the studies considered. In particular, patients included in the work by Castro-Villegas et al.[21] were more frequently treated with GCs (64.7% *versus* 30.0%), whereas those in the cohort used by Krintel et al.[20], showed a higher median CRP (15 *versus* 6 mg/l), TJC (15 *versus* 7), SJC (10 *versus* 1) and VAS-GH (70 *versus* 60) as compared to our cohort. In both previous studies the prediction model with miRNAs was not compared to or combined with baseline clinical parameters, which would have made the predictive estimations more generalizable. Concluding, heterogeneity cannot explain why results could not be validated within our study, although it might explain to a certain extent why the predictors identified across the studies are different.

Another possible contributing factor is represented by the chosen study design, in terms of inclusion criteria, measurement and time point of response, statistical analyses etc. Krintel et al. [20], analyzed a cohort of TNFi-naïve patients treated with MTX and intra-articular triamcinolone, which were additionally randomized to ADA treatment (n=90) or placebo treatment (n=90). To identify miRNAs specifically predictive for EULAR response to ADA combination therapy, an interaction term for each miRNA with the received treatment was added to the prediction model. Castro-Villegas et al. [21] used a cohort of patients treated with ADA, ETN or infliximab and investigated how serum miRNAs changed over time in ten patients after TNFi initiation using miScript miRNA PCR array (Qiagen, Hilden, Germany). In the following step, the ten most relevant miRNAs were measured by single-miRNA assays in 85 additional patients, and univariate and multivariable tests were applied to predict response to any TNFi. In case the true predictive ability of miRNAs is weak, the design might make the difference in inclusion or exclusion of each miRNA, which would explain the differences in identified miRNAs across studies. In addition, discrepancy in found miRNAs across studies might arise from the fact that each TNFi treatment is analyzed independently (our study and Krintel et al. [19]) or in combination with others to find universal miRNAs for TNFi response (Castro-Villegas et al. [20]). Indeed,

despite that all registered TNF-alpha inhibiting therapies target TNF-alpha, they have small chemical differences and etanercept, in particular, also targets lymphotoxin-alpha[32]. It is therefore possible that biomarkers predictive of response to TNFi-therapy are to some extent TNFi-specific.

Despite the advantage in terms of stability, the identification of circulating miRNAs with a concrete potential of application in clinical practice is very limited. In other inflammatory diseases, such as inflammatory bowel disease, the usage of miRNAs as potential biomarkers is still being explored, though has so far not revealed usable predictors of response to therapy [33]. In this line, the usefulness and robustness of miRNAs as biomarkers has been questioned; e.g. in non-neoplastic diseases only 33% (139/416) of the reported miRNAs were considered either biologically plausible, specific for the disease or interpretable with the current knowledge [34]. Another study showed that up to 58% of the reported circulating miRNAs related to cancer-subtypes were not disease specific and most likely derived from blood cells [35]. These studies indicate that false positive results in studies exploring circulating miRNA are lurking, and warrant additional carefulness when proposing a miRNA as a marker for a specific disease or disease state. For RA specifically, biomarkers are frequently identified in the circulation [13-19]. However, because the disease primarily affects the joints, the synovial compartment might constitute an alternative good source of biomarkers for RA, as has been demonstrated for cell-derived microparticles in the synovium compared to the circulation [36-38]. Since in RA the synovial miRNAs do not necessarily correlate with plasma miRNAs [12], it is possible that synovial miRNAs have better predictive abilities than the circulating ones. If so, incorporation of any predictive test on the synovial fluid or synovial tissue instead of the circulation, will affect the clinical feasibility negatively. In the future, the discovery of potential biomarker can be boosted by the implementation of novel high throughput techniques. One of the most promising at this regard is next-generation sequencing (NGS) that has the potential to also identify novel and not previously annotated miRNAs (currently 1882 known (<http://www.mirbase.org>, accessed July 4th, 2016 [39]), as it is not restricted to a predefined selection of miRNAs such as the multiplex based techniques

used in this study [40]. NGS might identify previously unknown targets and discover novel miRNAs for the prediction of response to RA.

## **CONCLUSION**

So far, there are no miRNAs that can be used in the prediction of response to TNFi therapy. We believe that a combination of differences in study design, technical variability, lack of multiple testing corrections, and heterogeneity between studies could contribute to these discrepancies. However, it is also conceivable that the irreproducibility of results is caused by the absence of truly biologically relevant miRNAs in the prediction of response to TNFi. Overall our study demonstrated that in order to increase reproducibility of the results, future studies addressing this topic should (1) standardize detection methods, (2) investigate the added value of miRNAs over clinical parameters, (3) technically replicate findings using a method suitable in case of implementation in clinical practice (i.e. single assay), (4) validate findings in a separate cohort, especially when correction for multiple testing in the discovery phase is not performed, and (5) since heterogeneity influences the ORs of miRNAs, prediction models should preferably be validated in a cohort close to the target population.

## REFERENCES

1. Wolfe AM, Kellgren JH, Masi AT. The epidemiology of rheumatoid arthritis: a review. II. Incidence and diagnostic criteria. *Bull Rheum Dis* 1968;19 3: 524-9.
2. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res* 2002;4 Suppl 3: S265-S72.
3. Keystone EC, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum* 2004;50 5: 1400-11.
4. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med* 1999;340 4: 253-9.
5. Cuppen BV, Welsing PM, Sprengers JJ, Bijlsma JW, Marijnissen AC, van Laar JM et al. Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability. *Rheumatology (Oxford)* 2016;55 5: 826-39.
6. Mookherjee N, El-Gabalawy HS. High degree of correlation between whole blood and PBMC expression levels of miR-155 and miR-146a in healthy controls and rheumatoid arthritis patients. *J Immunol Methods* 2013;400-401: 106-10.
7. Sayed D, Abdellatif M. MicroRNAs in development and disease. *Physiol Rev* 2011;91 3: 827-87.
8. Long L, Yu P, Liu Y, Wang S, Li R, Shi J et al. Upregulated microRNA-155 expression in peripheral blood mononuclear cells and fibroblast-like synoviocytes in rheumatoid arthritis. *Clin Dev Immunol* 2013;2013: 296139.
9. Wang J, Chen J, Sen S. MicroRNA as Biomarkers and Diagnostics. *J Cell Physiol* 2016;231 1: 25-30.
10. Cortez MA, Bueso-Ramos C, Ferdin J, Lopez-Berestein G, Sood AK, Calin GA. MicroRNAs in body fluids--the mix of hormones and biomarkers. *Nat Rev Clin Oncol* 2011;8 8: 467-77.
11. Pauley KM, Cha S, Chan EK. MicroRNA in autoimmunity and autoimmune diseases. *J Autoimmun* 2009;32 3-4: 189-94.
12. Murata K, Yoshitomi H, Tanida S, Ishikawa M, Nishitani K, Ito H et al. Plasma and synovial fluid microRNAs as potential biomarkers of rheumatoid arthritis and osteoarthritis. *Arthritis Res Ther* 2010;12 3: R86.
13. Duroux-Richard I, Jorgensen C, Apparailly F. What do microRNAs mean for rheumatoid arthritis? *Arthritis Rheum* 2012;64 1: 11-20.
14. Luo X, Ranade K, Talker R, Jallal B, Shen N, Yao Y. microRNA-mediated regulation of innate immune response in rheumatic diseases. *Arthritis Res Ther* 2013;15 2: 210.
15. Wittmann J, Jack HM. microRNAs in rheumatoid arthritis: midget RNAs with a giant impact. *Ann Rheum Dis* 2011;70 Suppl 1: i92-i6.
16. Bottini N, Firestein GS. Epigenetics in rheumatoid arthritis: a primer for rheumatologists. *Curr Rheumatol Rep* 2013;15 11: 372.
17. Ceribelli A, Yao B, Dominguez-Gutierrez PR, Nahid MA, Satoh M, Chan EK. MicroRNAs in systemic rheumatic diseases. *Arthritis Res Ther* 2011;13 4: 229.
18. Salehi E, Eftekhari R, Oraei M, Gharib A, Bidad K. MicroRNAs in rheumatoid arthritis. *Clin Rheumatol* 2015;34 4: 615-28.
19. Churov AV, Oleinik EK, Knip M. MicroRNAs in rheumatoid arthritis: altered expression and diagnostic potential. *Autoimmun Rev* 2015;14 11: 1029-37.
20. Krintel SB, Dehlendorff C, Hetland ML, Horslev-Petersen K, Andersen KK, Junker P et al. Prediction of treatment response to adalimumab: a double-blind placebo-controlled study of circulating microRNA in patients with early rheumatoid arthritis. *Pharmacogenomics J* 2015;16 2: 141-6.
21. Castro-Villegas C, Perez-Sanchez C, Escudero A, Filipescu I, Verdu M, Ruiz-Limon P et al. Circulating miRNAs as potential biomarkers of therapy effectiveness in rheumatoid arthritis patients treated with anti-TNFalpha. *Arthritis Res Ther* 2015;17: 49.
22. Duroux-Richard I, Pers YM, Fabre S, Ammari M, Baeten D, Cartron G et al. Circulating miRNA-125b is a potential biomarker predicting response to rituximab in rheumatoid arthritis. *Mediators Inflamm* 2014;2014: 342524.

23. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38 1: 44-8.
24. van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;41 10: 1845-50.
25. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001;25 4: 402-8.
26. Youden W, . Index for rating diagnostic tests. *Cancer* 1950;3 1: 32-5.
27. Ma R, Jiang T, Kang X. Circulating microRNAs in cancer: origin, function and application. *J Exp Clin Cancer Res* 2012;31: 38.
28. Witwer KW. Circulating microRNA biomarker studies: pitfalls and potential solutions. *Clin Chem* 2015;61 1: 56-63.
29. Farr RJ, Januszewski AS, Joglekar MV, Liang H, McAulley AK, Hewitt AW et al. A comparative analysis of high-throughput platforms for validation of a circulating microRNA signature in diabetic retinopathy. *Sci Rep* 2015;5: 10375.
30. Fujioka S, Nakamichi I, Esaki M, Asano K, Matsumoto T, Kitazono T. Serum microRNA levels in patients with Crohn's disease during induction therapy by infliximab. *J Gastroenterol Hepatol* 2014;29 6: 1207-14.
31. Nadal E, Truini A, Nakata A, Lin J, Reddy RM, Chang AC et al. A Novel Serum 4-microRNA Signature for Lung Cancer Detection. *Sci Rep* 2015;5: 12464.
32. Tansey MG, Szymkowski DE. The TNF superfamily in 2009: new pathways, new indications, and new drugs. *Drug Discov Today* 2009;14 23-24: 1082-8.
33. Kalla R, Ventham NT, Kennedy NA. MicroRNAs: new players in inflammatory bowel disease. *Gut* 2015;64 6: 1008.
34. Haider BA, Baras AS, McCall MN, Hertel JA, Cornish TC, Halushka MK. A critical evaluation of microRNA biomarkers in non-neoplastic disease. *PLoS One* 2014;9 2: e89565.
35. Pritchard CC, Kroh E, Wood B, Arroyo JD, Dougherty KJ, Miyaji MM et al. Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. *Cancer Prev Res (Phila)* 2012;5 3: 492-7.
36. Knijff-Dutmer EA, Koerts J, Nieuwland R, Kalsbeek-Batenburg EM, van de Laar MA. Elevated levels of platelet microparticles are associated with disease activity in rheumatoid arthritis. *Arthritis Rheum* 2002;46 6: 1498-503.
37. Vinuela-Berni V, Doniz-Padilla L, Figueroa-Vega N, Portillo-Salazar H, Abud-Mendoza C, Baranda L et al. Proportions of several types of plasma and urine microparticles are increased in patients with rheumatoid arthritis with active disease. *Clin Exp Immunol* 2015;180 3: 442-51.
38. Berckmans RJ, Nieuwland R, Kraan MC, Schaap MC, Pots D, Smeets TJ et al. Synovial microparticles from arthritic patients modulate chemokine and cytokine release by synoviocytes. *Arthritis Res Ther* 2005;7 3: R536-R44.
39. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 2014;42 Database issue: D68-D73.
40. Pimentel F, Bonilla P, Ravishankar YG, Contag A, Gopal N, LaCour S et al. Technology in MicroRNA Profiling: Circulating MicroRNAs as Noninvasive Cancer Biomarkers in Breast Cancer. *J Lab Autom* 2015;20 5: 574-88.

**Supplementary Table 1:** Baseline characteristics of responders and non-responders split for cohort and treatment. Patients were selected from the observational BiOCURA cohort based on treatment outcome over the course of one year after start of either ADA or ETN treatment. The clinical characteristics reported refer to the values present before treatment initiation. P-values for comparisons between responders and non-responders were calculated by means of an independent sample t-test, Mann-Whitney U test, fisher exact test (2\*2) or chi-square (>2\*2) based on the distribution of the clinical parameter. Bold p-values indicate significant associations (p<0.05).

### 1a Discovery cohort

Item	ADA (n=40)		p-value	ETN (n=40)		p-value
	Non-resp (n=20)	Resp (n=20)		Non-resp (n=20)	Resp (n=20)	
Female gender, n (%)	15 (75)	14 (70)	1.00	17 (85)	15 (75)	0.70
Age, mean years $\pm$ sd	53.4 $\pm$ 11.9	56.4 $\pm$ 12.9	0.46	56.3 $\pm$ 9.7	53.9 $\pm$ 10.1	0.44
Current smoker, n (%)	8 (40)	5 (25)	0.50	7 (35)	4 (20)	0.48
RF positivity, n (%)	11 (55)	16 (80)	0.18	14 (70)	15 (75)	0.72
ACPA positivity, n (%)	12 (60)	15 (75)	0.50	14 (70)	17 (85)	0.13
CRP, median	10.0 (4.0-13.0)	7.5 (4.0-12.8)	0.97	4.3 (3.1-7.5)	8.5 (3.8-30.0)	0.07
No. of prev. bDMARDs			0.14			0.57
0, n (%)	13 (65)	17 (85)		15 (75)	15 (75)	
1, n (%)	7 (35)	3 (15)		4 (20)	5 (25)	
2, n (%)	0 (0)	0 (0)		1 (5)	0 (0)	
Concomitant treatment, n (%)	19 (95)	19 (95)	1.00	20 (100)	19 (95)	1.00
MTX, n (%)	13 (65)	17 (85)	0.27	13 (65)	16 (80)	0.48
SSZ, n (%)	2 (10)	3 (15)	1.00	3 (15)	1 (5)	0.61
HCQ (%)	5 (25)	5 (25)	1.00	7 (35)	6 (30)	1.00
GCs, n (%)	8 (40)	3 (15)	0.16	7 (35)	5 (25)	0.73
Baseline DAS28, mean $\pm$ sd	4.3 $\pm$ 1.3	4.8 $\pm$ 1.0	0.15	4.4 $\pm$ 1.2	4.8 $\pm$ 1.0	0.28
TJC, median (IQR)	7.0 (3.0-13.8)	7.0 (4.3-14.5)	0.88	4.5 (1.0-21.5)	6.5 (2.0-17.8)	1.00
SJC, median (IQR)	1.0 (0.0-2.8)	3.0 (0.0-6.8)	0.08	1.0 (0.0-3.0)	1.5 (0.3-4.8)	0.30
VAS, mean $\pm$ sd	62.0 $\pm$ 20.7	64.5 $\pm$ 23.8	0.73	54.5 $\pm$ 26.0	54.5 $\pm$ 23.9	1.00
ESR, median mm/hr (IQR)	12.5 (2.5-30.0)	18.5 (9.0-29.8)	0.33	20.0 (5.3-35.3)	27.0 (13.0-54.3)	0.23

### 1b Validation cohort

Item	ADA (n=20)			ETN (n=20)		
	Non-resp (n=10)	Resp (n=10)	p-value	Non-resp (n=10)	Resp (n=10)	p-value
Female gender, n (%)	6 (60)	7 (70)	1.00	8 (80)	6 (60)	0.63
Age, mean years $\pm$ sd	56.4 $\pm$ 8.9	47.6 $\pm$ 10.8	0.06	62.3 $\pm$ 6.8	57.7 $\pm$ 11.5	0.28
Current smoker, n (%)	7 (70)	4 (40)	0.37	1 (10)	3 (30)	0.58
RF positivity, n (%)	5 (50)	5 (50)	1.00	6 (60)	6 (60)	1.00
ACPA positivity, n (%)	7 (70)	4 (40)	0.37	5 (50)	8 (80)	0.35
CRP, median	1.6 (1.0-6.8)	2.7 (1.0-10.8)	0.80	2.0 (1.0-23.5)	8.0 (3.3-12.3)	0.53
No. of prev. bDMARDs			0.42			0.55
0, n (%)	7 (70)	6 (60)		7 (70)	7 (70)	
1, n (%)	2 (20)	4 (40)		3 (30)	2 (20)	
2, n (%)	1 (10)	0 (0)		0 (0)	1 (10)	
Concomitant treatment, n (%)	10 (100)	10 (100)	-	7 (70)	10 (100)	0.21
MTX, n (%)	8 (80)	10 (100)	0.47	5 (50)	9 (90)	0.14
SSZ, n (%)	0 (0)	1 (10)	1.00	1 (10)	1 (10)	1.00
HCQ (%)	3 (30)	2 (20)	1.00	3 (30)	5 (50)	0.65
GCs, n (%)	6 (60)	1 (10)	0.06	4 (40)	1 (10)	0.30
Baseline DAS28, mean $\pm$ sd	3.2 $\pm$ 1.3	4.5 $\pm$ 0.7	<b>0.02</b>	4.1 $\pm$ 1.2	4.5 $\pm$ 0.8	0.55
TJC, median (IQR)	2.0 (0.0-11.5)	5.5 (3.5-14.3)	0.12	7.0 (2.8-15.0)	4.0 (3.0-10.0)	0.63
SJC, median (IQR)	0.0 (0.0-0.0)	0.5 (0.0-3.0)	0.12	1.0 (0.0-4.0)	2.0 (0.8-3.0)	0.48
VAS, mean $\pm$ sd	41.5 $\pm$ 24.7	62.5 $\pm$ 18.7	0.05	57.5 $\pm$ 15.9	52.0 $\pm$ 23.0	0.54
ESR, median mm/hr (IQR)	9.0 (4.5-21.3)	14.0 (9.3-38.8)	0.25	9.5 (5.3-22.5)	18.5 (13.0-31.5)	0.17

ACPA, anti-citrullinated protein antibody; ADA, adalimumab; bDMARDs, biological disease modifying antirheumatic drugs; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; ETN, etanercept; GC, glucocorticoid; HCQ, hydroxychloroquine; IQR, interquartile range; MTX, methotrexate; RF, rheumatoid factor; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count; VAS-GH, visual analogue scale general health.

**Supplementary Table 2:** Baseline characteristics of patients in the discovery and validation cohort, split for received treatment. Patients were selected from the observational BiOCURA cohort based on treatment outcome over the course of one year after start of either ADA or ETN treatment. The presented clinical characteristics refer to the values present before treatment initiation. P-values for comparisons were calculated by means of an independent sample t-test, Mann-Whitney U test, fisher exact test (2\*2) or chi-square (>2\*2) based on the distribution of the clinical parameter. Bold p-values indicate significant associations (p<0.05).

Item	Adalimumab			Etanercept		
	Discovery (n=40)	Validation (n=20)	p-value	Discovery (n=40)	Validation (n=20)	p-value
Female gender, n (%)	29 (72.5)	13 (65.0)	0.56	32 (80.0)	14 (70.0)	0.52
Age, mean years ±sd	54.9 ±12.3	52.0 ±10.6	0.37	55.1 ±9.8	60.0 ±9.5	0.07
Current smoker, n (%)	13 (32.5)	11 (55.0)	0.11	11 (27.5)	4 (20.0)	0.75
RF positivity, n (%)	27 (67.5)	10 (50.0)	0.26	30 (75.0)	12 (60.0)	0.25
ACPA positivity, n (%)	27 (67.5)	11 (55.0)	0.40	32 (80)	13 (65.0)	0.22
CRP, median (IQR)	8.0 (4.0-13.0)	1.7 (1.0-7.5)	<b>&lt;0.01</b>	6.0 (3.2-12.0)	5.5 (1.3-12.8)	0.29
No. of previously used bDMARDs			0.32			0.85
0, n (%)	30 (75.0)	13 (65.0)		30 (75.0)	14 (70.0)	
1, n (%)	10 (25.0)	6 (30.0)		9 (22.5)	5 (25.0)	
2, n (%)	0 (0.0)	1 (5.0)		1 (2.5)	1 (5.0)	
Concomitant treatment, n (%)	38 (95.0)	20 (100.0)	0.55	39 (97.5)	17 (85.0)	0.10
Methotrexate, n (%)	30 (75.0)	18 (90.0)	0.30	29 (72.5)	14 (70.0)	1.00
Sulfasalazine, n (%)	5 (12.5)	1 (5.0)	0.65	4 (10.0)	2 (10.0)	1.00
Hydroxychloroquin, n (%)	10 (25.0)	5 (25.0)	1.00	13 (32.5)	8 (40.0)	0.58
Glucocorticoids, n (%)	12 (30.0)	7 (35)	0.72	12 (30.0)	5 (25.0)	0.77
Baseline DAS28, mean ±sd	4.6 ±1.2	3.8 ±1.2	<b>0.03</b>	4.6 ±1.1	4.3 ±1.0	0.40
TJC, median (IQR)	7.0 (3.0-13.8)	4.5 (1.0-12.5)	0.15	5.5 (1.3-14.8)	6.0 (3.0-13.3)	0.75
SJC, median (IQR)	2.0 (0.0-4.0)	0.0 (0.0-1.8)	<b>&lt;0.01</b>	1.0 (0.0-4.0)	1.0 (0.0-3.0)	0.72
VAS, mean ±sd	63.3 ±22.1	52.0 ±23.9	0.08	54.5 ±24.6	54.8 ±19.4	0.97
ESR, median mm/hr (IQR)	13.0 (6.3-29.8)	12.5 (6.3-30.3)	0.86	21.5 (11.3-40.8)	14.0 (7.3-30.8)	0.09

ACPA, anti-citrullinated protein antibody; ADA, adalimumab; bDMARDs, biological disease modifying antirheumatic drugs; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; ETN, etanercept; GC, glucocorticoid; HCQ, hydroxychloroquine; IQR, interquartile range; MTX, methotrexate; RF, rheumatoid factor; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count; VAS-GH, visual analogue scale general health.

**Supplementary Table 3:** Validation of multivariable models for prediction of response to TNFi. The prediction models from the discovery cohort, based on clinical parameters alone and miRNAs, were applied in the validation cohort (n=20 ADA or n=20 ETN) including the miRNA levels as measured by single assays. For ADA, the clinical baseline characteristics predicting response were (the square root of) the SJC, DAS28, GC use; the combined model included also miR99a and miR197. For ETN, the clinical model contained the (logtransformed) CRP only and the combined model also included miR-23a and miR-143. In order to limit re-fitting, the prediction rules of the clinical and combined models were applied on the validation cohort while freezing the regression coefficients of the individual parameters in the model. Per model, the area under the AUC-ROC was calculated and the sensitivity and specificity were shown for the best cut-off value, according to the Youden's index.

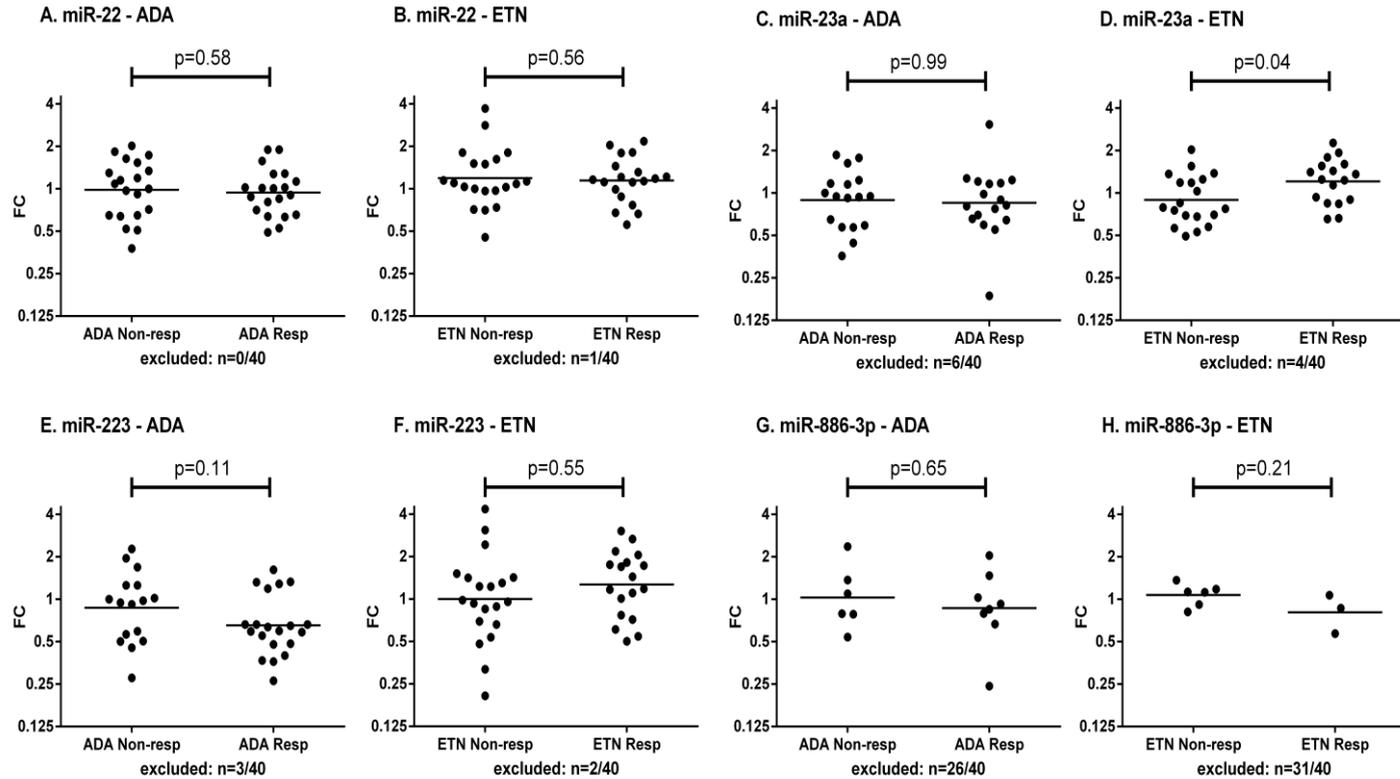
TNFi	Model	Model content	AUC-ROC	Sens.	Spec.
<b>ADA</b>	Clinical	SJC, GC use, DAS28	0.93	90%	100%
	Clinical + miRNAs	SJC, GC use, DAS28, miR-99a, miR-143	0.57	50%	80%
<b>ETN</b>	Clinical	CRP	0.59	80%	70%
	Clinical + miRNAs	CRP, miR-23a, miR-197	0.66	80%	60%

ADA, adalimumab; AUC-ROC, area under the receiver operating characteristic curve CRP, C-reactive protein; ETN, etanercept; GC, glucocorticoid; sens, sensitivity; SJC, swollen joint count; spec, specificity; TNFi, TNF- $\alpha$ -inhibitor.

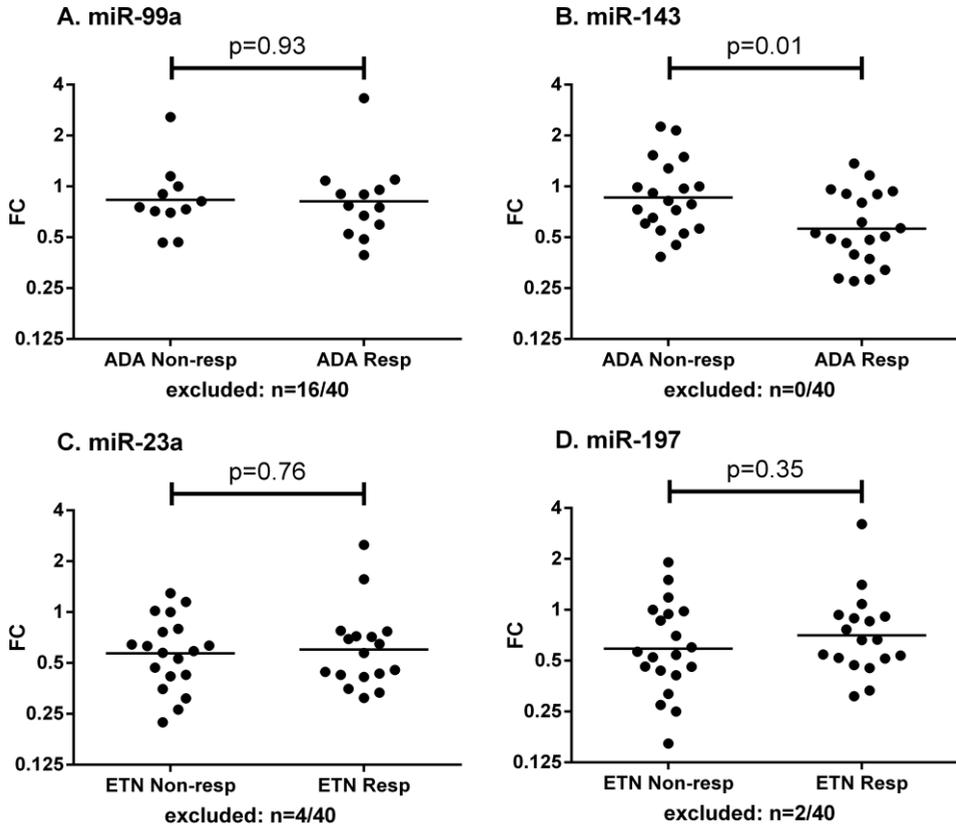
**Supplementary Table 4:** Clinical parameters at baseline associated with current miRNA levels. Four miRNAs were measured using Taqman single assays in 80 patient from discovery cohort (i.e. technical replication), and 40 from validation. To evaluate which clinical parameters were associated with current miRNA values, per miRNA a multivariable linear regression model was built with the miRNA level ( $-\Delta\Delta Ct$ ) as dependent variable and baseline clinical parameters as independent variables (full model). Clinical parameters that were entered in the full model were gender, age, smoking status, methotrexate use, hydroxychloroquine use, glucocorticoid use, DAS28, swollen joint count (SJC, square root), erythrocyte sedimentation rate (ESR, natural logarithm) and C-reactive protein (CRP, natural logarithm). Subsequently, a backward selection procedure ( $p$ -value $>0.05$  for exclusion) was used to come to the definitive selection of most influential parameters (final model). Per miRNA, the clinical parameters in the final models were shown that significantly associated with current miRNA levels. The R-square value of each model was determined, as an indication of the explained variance, i.e. the proportion of all variability in miRNA levels that can be explained by the clinical parameters in the model. We did not separate for ADA or ETN treatment received, since all measurements in this analysis were collected before treatment initiation.

miRNA	Clinical parameter	Effect on miRNA level	p-value	R-square of model
<b>miR99a</b>	CRP	Increase	<0.01	0.14
<b>miR143</b>	Age	Decrease	<0.01	0.34
	CRP	Increase	0.02	
	ESR	Decrease	<0.01	
<b>miR23a</b>	Gender, female	Decrease	0.02	0.14
	ESR	Decrease	0.02	
<b>miR197</b>	CRP	Decrease	<0.01	0.15
	ESR	Increase	<0.01	

**Supplementary Figure 1:** Expression of miRNAs predicting TNFi response identified in other studies. In previous studies[20, 21] baseline values of miR-22 (A & B), miR-23a (C & D), miR-223 (E & F) and miR-886-3p (G & H) were found to be associated with response to TNFi treatment. The expression (FC of individual patients and geometric mean) of these miRNAs are shown in the discovery cohort (n=40 ADA and n=40 ETN) as measured by the OpenArray platform. P-values were calculated on the  $-\Delta\Delta C_{rt}$  values using an independent sample t-test.



**Supplementary Figure 2:** Technical replication of selected miRNAs from the discovery cohort. Samples from the discovery cohort were re-analyzed using single miRNA assays for the selected miRNAs. If samples were excluded from the discovery step because of low amplification scores, these were also excluded from the technical replication. Graphs for ADA (A & B) and ETN (C & D) report the levels of expression as fold change (FC) per individual patient and the geometric means per group (FC). P-values were calculated on the  $-\Delta\Delta\text{Ct}$  values using an independent sample t-test.







## Chapter 5

### **RNA-Sequencing to predict response to TNF- $\alpha$ inhibitors reveals possible mechanism for non-response in smokers**

Bart V.J. Cuppen<sup>#</sup>, Marzia Rossato<sup>#</sup>, Ruth D.E. Fritsch-Stork, Arno N. Concepcion, Suzanne Linn-Rasker, Johannes W.J. Bijlsma<sup>1</sup>, Jacob M. van Laar, Floris P.J.G. Lafeber, Timothy R. Radstake, on behalf of all SRU investigators

# These authors contributed equally

Submitted

## **ABSTRACT**

### **Objective**

Many studies have employed messenger RNA (mRNA) profiling to predict response to tumor necrosis factor-alpha inhibitors (TNFi) in rheumatoid arthritis (RA), yet efforts to validate these targets have failed to show predictive abilities acceptable for clinical practice. We explored the transcriptome to identify and validate markers for response to two TNFi therapies: adalimumab (ADA) and etanercept (ETN).

### **Methods**

The most extreme responders and non-responders to TNFi therapy were selected from the observational BiOCURA cohort and RNA-sequencing was performed on mRNA derived from peripheral blood mononuclear cells (PBMCs) collected before initiation of treatment. The relative up/downregulation of pathways as well as individual gene transcripts in responders and non-responders was investigated. Promising targets were technically replicated and validated in a different set of patients using quantitative PCR assays.

### **Results**

In non-responders, pathways related to interferon and cytokine signaling were relatively downregulated at baseline, whereas pathways activated in response to the influenza virus were upregulated. Additionally, among 178 genes significantly differentially expressed between responders and non-responders, GPR15 and SEMA6B were considered the most interesting targets predictive for response to ADA and ETN. Both of these targets were (to some extent) validated in a separate cohort of patients treated with ADA, while they were not predictive of response in a new set of patients treated with ETN. Additional analyses revealed that GPR15 and SEMA6B did not independently predict response, but were rather dose-dependent markers for smoking.

### **Conclusion**

This study was not able to identify new transcripts ready to use in clinical practice, yet GPR15 and SEMA6B were identified as candidate explanatory markers for the reduced treatment success observed in RA smokers.

## INTRODUCTION

Rheumatoid arthritis (RA) is a chronic, disabling disease that mainly affects the synovial joints. With a large arsenal of treatments among which tumor necrosis factor-alpha inhibitors (TNFi's), RA disease activity decreases sufficiently in most but not all cases. The identification of responders and non-responders before initiation of therapy would therefore aid in making strategic treatment decisions and improve clinical outcomes. Thus far, however, no biomarker to predict response to TNFi treatment is robust enough to be used in clinical practice[1].

Messenger RNAs (mRNAs) are of potential interest as biomarkers for response since they constitute the bridge between the genetic and protein formation. In order to identify mRNA gene transcripts able to predict response to biological treatment, the transcriptome has been studied extensively by microarrays [2, 3]. However, the results generated by these studies are heterogeneous and efforts to replicate models in external cohorts failed to reach predictive abilities acceptable for clinical practice [4-6]. The inability to validate results in multiple cohorts might be partially related to the technical drawbacks of microarrays. RNA-sequencing (RNA-seq) is a newer technique that overcomes multiple limitations characterizing microarrays and therefore eventually used more routinely for the measurement of mRNA levels [7, 8]. Whereas microarray profiling is a chip-based method using a selected set of gene transcripts, RNA-seq is not based on a predefined platform, and it is thus able to identify expressed genes that have not previously been annotated [9]. RNA sequencing is also able to quantify genes with very low and high expression more accurately, because it is characterized by less background noise and has no (cross) hybridization issues [9, 10]. In addition, thanks to its bigger dynamic range, RNA sequencing is more suitable for the quantification of absolute gene expression, thus resulting in more reproducible data [11-13]. One recent study employed RNA-sequencing of neutrophils to predict the response to TNFi's and found that a combination of interferon regulated transcripts at baseline predicted the response to TNFi's [14]. Carefulness with the validity of these results is warranted as no attempts were undertaken to replicate these results in another cohort of patients.

In this study we continued the search for gene transcripts predicting therapeutic response, by employing RNA-sequencing on baseline mRNA of peripheral blood mononuclear cells

(PBMCs) in RA patients treated with two distinct TNFi treatments, namely adalimumab (ADA) and etanercept (ETN). Additionally, we validated the identified predictors in a separate cohort and investigated the association of gene expression with PBMC cell subsets and clinical characteristics of patients.

## **METHODS**

### **Clinical data collection**

Patients initiating ADA or ETN therapy were selected from the “Biologicals and Outcome Compared and predicted Utrecht region in Rheumatoid Arthritis” (BiOCURA) study. BiOCURA is an observational cohort, in which RA patients eligible for biological treatment according to regular clinical practice were enrolled and followed after start of treatment, in one academic hospital and seven regional hospitals in the Netherlands. Re-inclusion after switching to a different biological treatment was possible, at which patients re-entered baseline again. The study was approved by the local ethics committee of the University Medical Center Utrecht and the institutional review boards of the participating centers (see Acknowledgements). Informed consent was obtained from each patient.

Trained nurses gathered all data during a dedicated visit, which included all clinical parameters, joint counts and collection of blood. Visits were scheduled at baseline (before initiation) and after three, six and twelve months of treatment. Disease activity was assessed using the disease activity score based on a 28-joint count (DAS28)[15] and subsequently the European League Against Rheumatism (EULAR) response compared with baseline was calculated [16]. This study design allowed the determination of a clinical response of each patient, on the basis of three responses over the course of one year.

### **Patient selection**

Two separate cohorts were composed: a discovery cohort to select potentially predictive gene transcripts, and a validation cohort to test if the results found in the discovery phase were reproducible. The discovery cohort was formed by selecting the most extreme patients regarding clinical response, among all ADA and ETN treated patients included

between June 2009 (start of BiOCURA) and October 2012 (n=74 ADA and n=68 ETN). The top responding patients (n=20 for both ADA and ETN, from now on called “responders”), were identified by the selection of patients with the best three EULAR responses over the course of one year. The selection of poor responders (n=20 for both ADA and ETN, from now on called “non-responders”), was based on the least optimal EULAR responses over the course of one year and/or (early) discontinuation of TNFi treatment due to inefficacy. Among the 20 responders and non-responders to ADA and ETN, ten patients with most extreme responses were identified to analyze groups of patients with even clearer treatment outcomes (from now on called “extreme responders” and “extreme non-responders”). Patients with a baseline DAS28 <2.6, usually due to involvement of joints that are not calculated in the 28-joint count such as in the foot, were excluded from the analysis in order to reduce the chance that limited improvement in DAS28 resulted in misclassification as EULAR non-responders. For the validation cohort, responders (n=10 for both ADA and ETN each) and non-responders (n=10 for both ADA and ETN each) were selected using the same criteria as in the discovery cohort, among patients included from October 2012 until June 2015 (n=25 ADA and n=40 ETN). Due to the smaller size of the cohort from which patients were selected, the differences in clinical outcome between responders and non-responders in the validation cohort was less extreme. The baseline characteristics for responders and non-responders are shown in Table 1 and for responders and non-responders split per cohort in Supplementary Table 1.

**Table 1:** Baseline characteristics of patients, split for treatment and response. Patients were selected from the observational BiOCURA cohort based on treatment outcome over the course of one year after start of either ADA or ETN. The presented clinical characteristics for responders and non-responders are all before treatment initiation. P-values for comparisons were calculated by means of an independent sample t-test, Mann-Whitney U test, fisher exact test (2\*2) or chi-square (>2\*2) based on distribution of the clinical parameter.

Item	ADA (n=60)		p-value	ETN (n=60)		p-value
	Non-resp (n=30)	Resp (n=30)		Non-resp (n=30)	Resp (n=30)	
Female gender, n (%)	21 (70)	21 (70)	1.00	25 (83)	21 (70)	0.36
Age, mean years $\pm$ sd	54.4 $\pm$ 10.9	53.5 $\pm$ 12.7	0.76	58.3 $\pm$ 9.2	55.1 $\pm$ 10.5	0.22
Current smoker, n (%)	16 (53)	8 (26.7)	0.06	8 (26.7)	7 (23.3)	1.00
RF positivity, n (%)	16 (53)	21 (70)	0.29	20 (67)	22 (73)	0.78
ACPA positivity, n (%)	19 (63)	19 (63)	1.00	19 (63)	26 (87)	0.07
CRP, median (IQR)	5.2 (1.6-10.5)	5.5 (2.0-12.3)	0.78	4.0 (2.0-9.0)	8.5 (4.0-18.3)	<b>0.03</b>
No. of previously used bDMARDs			0.48			1.00
0	20 (67)	23 (78)		22 (73)	22 (73)	
1	9 (30)	7 (23)		7 (23)	7 (23)	
2	1 (3)	0 (0)		1 (3)	1 (3)	
Concomitant treatment, n (%)	29 (97)	29 (97)	1.00	27 (90)	29 (97)	0.61
MTX, n (%)	21 (70)	27 (90)	0.10	18 (60)	25 (83)	0.08
SSZ, n (%)	2 (7)	4 (13)	0.67	4 (13)	2 (7)	0.67
HCQ, n (%)	8 (27)	7 (23)	1.00	10 (33)	11 (37)	1.00
GC, n (%)	15 (50)	4 (13)	0.01	11 (37)	6 (20)	0.25
Baseline DAS28, mean $\pm$ sd	3.9 $\pm$ 1.4	4.7 $\pm$ 0.9	0.01	4.3 $\pm$ 1.2	4.6 $\pm$ 0.9	0.21
TJC, median (IQR)	5.0 (1.0-13.0)	7.0 (4.0-14.3)	0.35	6.5 (2.8-11.3)	5.0 (2.8-11.3)	0.87
SJC, median (IQR)	0.0 (0.0-4.0)	2.0 (0.0-4.0)	<b>0.03</b>	1.0 (0.0-3.3)	2.0 (0.8-4.0)	0.20
VAS-GH, mean $\pm$ sd	55.2 $\pm$ 23.8	63.8 $\pm$ 22.0	0.15	55.5 $\pm$ 22.8	55.1 $\pm$ 10.5	0.76
ESR, median (IQR)	11.0 (3.8-26.0)	16.5 (9.0-32.0)	0.14	13.0 (5.8-33.8)	21.0 (14.3-39.5)	0.07

ACPA, anti-citrullinated protein antibody; ADA, adalimumab; bDMARDs, biological disease modifying antirheumatic drugs; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; ETN, etanercept; GC, glucocorticoid; HCQ, hydroxychloroquine; IQR, interquartile range; MTX, methotrexate; RF, rheumatoid factor; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count; VAS-GH, visual analogue scale general health.

## **Blood processing and RNA extraction**

Blood was collected in a 70 ml lithium heparin tube and PBMCs were isolated using Ficoll Paque Plus (GE Healthcare, Uppsala, Sweden).  $5 \times 10^6$  PBMCs were lysed for total RNA isolation using RNeasy kit following the manufacturer's instructions (Qiagen). Quantification of RNA and purity was assessed using Nanodrop (Thermo Fischer Scientific, Waltham, Massachusetts, USA) and the quality with Bioanalyzer (Agilent, Santa Clara, California, USA). Samples with a RNA-integrity score (RIN) lower than 8 were excluded from the analysis. RNA was stored per 2  $\mu$ g at  $-80^\circ\text{C}$  until use.

## **mRNA analyses**

### RNA-Sequencing

RNA-Sequencing of the discovery cohort was performed at the Beijing Genomics Institute (BGI, Hong Kong). RNA-seq library preparation was performed starting from 100ng total RNA using the TruSeq kit (Illumina). Briefly, messenger RNA molecules were enriched by using the oligo(dT) magnetic beads and fragmented into short fragments (about 200 bp). After conversion to double stranded cDNA using random hexamer-primer, end reparation and 3'-end single nucleotide A (adenine) addition were performed. Finally, sequencing adaptors were ligated to the fragments that were subsequently enriched by PCR amplification. Quality and quantity of RNA-seq libraries were determined using the Agilent 2100 Bioanalyzer and the ABI StepOnePlus Real-Time PCR System respectively. The library products were sequenced on an Illumina HiSeq<sup>TM</sup> 2000 sequencer using 50bp single-end reads, generating 20 million clean reads per sample. After quality filtering according to the BGI pipeline, reads were aligned to the GrCh38 reference human genome (Genome Reference consortium) and the H. sapiens transcriptome (Ensembl, version 78) using SOAPaligner/SOAP2 [17]. Summed exon read counts per gene were estimated using HTSeq-count [18]. Differential expression analysis was performed using the negative binomial distribution-based method implemented in DESeq on the summed exon read counts per gene [19]. Group comparisons were performed between six subgroups: non-responders *versus* responders to ADA, ETN and all TNFi (pooled ADA and ETN), and extreme non-responders *versus* extreme responders to ADA, ETN and all TNFi. A filter was applied to exclude all genes with a mean reads per kilobase per million mapped reads

(RPKM) lower than 0.5, as genes with very low expression levels are less reliably measured by RNA-seq[8], and therefore are less suitable as biomarkers[20].

### Single qPCR assays

For the technical replication and validation experiments, Taqman gene expression RT-qPCR kits for SEMA6B (Hs00220339\_m1), GPR15 (Hs00922903\_s1), ACTB (Hs99999903\_m1) and GAPDH (Hs99999905\_m1) were purchased from LifeTechnologies, ThermoFisher. ACTB and GAPDH were selected as reference genes since they were both highly expressed and stable between responders and non-responders ( $\log_2FC < 0.03$ ). In short, 0.3 $\mu$ g of PBMC-derived RNA was used as starting material for both the technical replication (n=80) and validation analysis (n=40). cDNA was synthesized from RNA using the iScript<sup>tm</sup> synthesis kit (BIO RAD), according to the manufacturer's instructions, while using the following thermal cycler conditions: 5 min, 25°C, 30 min 42°C, 5 min 85°C. Quantification of gene transcripts was performed in duplicate from 2 $\mu$ l 6-times diluted cDNA, TaqMan specific primers of the gene expression assay and TaqMan Fast Advance Master Mix in a final volume of 15 $\mu$ l, on a Quantstudio 12kflex Real-Time PCR system (LifeTechnologies) using the following thermal cycler conditions: 2 min, 50°C; 20 sec, 95°C; 45 cycles of 1 sec, 95°C; 20 sec, 60°C. Data were analyzed according to the comparative threshold cycle method [21], after normalization by the selected housekeeping genes. The fold change (FC) was calculated based on one randomly chosen sample that was set as 1 and analyzed on each qPCR plate, thus allowing proper comparison of data across different plates.

### **Flow cytometry**

After isolation from whole blood, PBMCs were stained by using specific monoclonal antibodies and analyzed by fluorescence-activated cell sorting (FACS) using a FACS-LSRII (BD) in order to determine the abundance of specific mononuclear cell subsets: anti-CD45 (FITC, Beckman Coulter) was used to identify leukocytes, anti-CD3 (FITC, Beckman Coulter) for T-cells, anti-CD4 (PECy5, Dako) for CD4 T-cells, anti-CD8 (PE, BD) for CD8 T-cells, anti-CD14 (PE, Beckman Coulter) for monocytes, anti-CD16+56 (PE, Beckman Coulter) for NK cells, anti-CD19 (FITC, BD) for B-cells. Per sample, 30.000 events were

registered, after which these data were further analyzed using FlowJo software (Tree star, Ashland, OR, USA).

## **Statistical analyses**

### Pathway enrichment analysis

First it was verified whether the baseline expression of genes downstream of TNF $\alpha$  and lymphotoxin-alpha (LT $\alpha$ ) were differentially expressed in responders *versus* non-responders. Intracellular pathways activated by the TNF $\alpha$  and LT $\alpha$  receptors were derived from the Kyoto Encyclopedia of Genes and Genomes (KEGG [22], ID 04668 + 04064 (<http://www.genome.jp/kegg/pathway.html>)); the expression of genes belonging to these pathways was retrieved from the differential gene expression analysis, without the application of a restriction based on minimal expression (i.e. RPKM>0.5).

Pathway enrichment analyses was performed on the complete set of expressed genes (i.e. with mean expression > 0.5 RPKM), using the online available tool Gene Set Enrichment Analyses (GSEA[23]). GSEA is a computational method that determines whether a predefined set of genes, e.g. a pathway, shows statistically significant and concordant differences between two biological states, in this case responders and non-responders. GSEA analysis was performed on the relative expression of genes in all TNFi responders *versus* non-responders using the REACTOME pathway database [24] as source of gene lists. Enriched pathways with a Benjamini & Hochberg adjusted p-value (B&H adj. p-val) lower than 0.20 were considered significantly enriched [23]. Traditional pathway enrichment analysis was performed on protein-coding genes identified as differentially expressed in responders *versus* non-responders in any of the six group comparisons (see 2.4.2. Selection of genes), via the web-portal Toppgene (<https://toppgene.cchmc.org/enrichment.jsp>, [25]). The top-10 enriched pathways and Gene Ontology (GO) terms linked to biological processes having the lowest B&H adj. p-values were retrieved.

### Selection of genes

In order to identify the most interesting individual targets in the context of TNFi response, we selected genes that significantly predicted response in multiple subgroups. Genes with

absolute  $\log_2(\text{FC}) > 0.58$  (i.e. a corresponding  $\text{FC} < 0.67$  or  $> 1.5$ ) and  $p\text{-value} < 0.05$  were considered significantly modulated and relevant in the context of TNFi response. These cut-offs were set low to not miss any important genes, whereas the validation phase was supposed to test if estimates are reliable. We considered genes that emerged in multiple comparisons as most promising targets, especially when seen in both non-responders *versus* responders and extreme non-responders *versus* extreme responders, as this can be considered a small internal validation. The genes of particular interest were plotted in GraphPad Prism software (GraphPad, Lo Jolla, CA, USA) as the reads RPKM of (excellent) non-responders *versus* (excellent) responders. Sensitivities and specificities (i.e. the proportion of respectively responders and non-responders to TNFi that are correctly identified as such) for genes were calculated on the  $\log(\text{RPKM})$ -values with the cut-off based on the Youden's index[26].

#### Influencing parameters

We investigated whether the expression of selected genes as measured by single qPCR was related to the relative subset composition of PBMCs as measured by flow cytometry. After setting a gate for leukocytes on the forward and side scatter (FSC/SSC)-plot, we determined the proportion of T-cells (CD3 pos, CD16+56 neg), Th-cells (CD4 pos, CD8 neg), cytotoxic T-cells (CD8 pos, CD4 neg), monocytes (CD14 pos, CD45 pos), B-cells (CD19 pos, CD8 neg), NK cells (CD16+56 pos, CD3 neg) and NK T-cells (CD16+56 pos, CD3 pos) present in the PBMC population. Plots for gene expression *versus* percentage of cells on all 120 patients (development + validation cohort) were made in Graphpad, and correlations and p-values were calculated by a Spearman's  $r$ .

We additionally investigated whether the gene expression as measured by qPCR in the complete cohort ( $n=120$ ) associated with the clinical parameters: gender, age, smoking status, concomitant use of methotrexate, hydroxychloroquine or glucocorticoids, DAS28, (log-transformed) erythrocyte sedimentation rate (ESR) or (log-transformed) C-reactive protein (CRP). We performed a linear regression on the  $-\text{ddCt}$  values of the genes with these parameters as independent variables and performed a backward selection procedure ( $p < 0.05$ ), so that the most explanatory parameters remained.

## RESULTS

### **Interferon-related pathways are downregulated at baseline in non-responders to TNFi**

In order to identify gene transcripts and pathways differentially expressed in responders *versus* non-responders before initiation of ADA and ETN treatment, transcriptome analysis by RNA-seq was performed on baseline PBMCs from a cohort of 80 RA patients. As ADA and ETN both target TNF $\alpha$ , a ligand for TNF-receptor I (TNF-RI) and TNF-RII, it is conceivable that any altered gene in the downstream pathway of this cytokine can affect the chances of response. Additionally, as ETN also targets LT $\alpha$  (ligand for TNF-RI, TNF-RII and LT $\beta$ -R) [27, 28], it is possible that a predictor for ADA is not predictive for ETN, and vice versa. Therefore, we first investigated whether the baseline expression of genes belonging to the pathways activated by TNF $\alpha$  and LT $\alpha$  were concomitantly up- or downregulated in responders to ADA and/or ETN (Supplementary Table 2). Of all genes belonging to these pathways, none was significantly differentially expressed in more than one subgroup, with the only exception being MAPK10 which was downregulated in ETN non-responders (FC=0.579, p=0.046) and in ETN extreme non-responders (FC=0.444, p=0.045) prior to treatment. However, MAPK10 showed extremely low expression (RPKM<0.5 in all samples), which questions the reliability of these findings and the likelihood of successful implementation into assays to predict individual TNFi response [8, 20].

Next, we investigated any pathways either up- or down-regulated in the entire set of measured transcripts with a minimal level of expression (RPKM>0.5, i.e. 12,716 genes), using Gene Set Enrichment Analysis (GSEA) analysis. In total, 26 pathways were significantly up- or downregulated in non-responders compared to responders prior to treatment. Among these, seven pathways could be directly linked to immune response: three interferon-related (all downregulated), two influenza life-cycle related (both upregulated), one cytokine signaling related (downregulated) and one MHC class II related (downregulated) (Supplementary Table 3). Further investigation on the gene content of these seven pathways (Supplementary file 1) demonstrated that around 67% (228 of 338) of all genes were annotated to multiple immunologic pathways. The most overlapping pathways with respect to their gene content were the three interferon related and the “cytokine signaling in immune system”, with a relative overlap in all six comparisons

ranging from 20% to 56%. The genes annotated to the two influenza life-cycle related pathways overlapped with each other for 75%, yet hardly with any other pathways, mainly due to the inclusion of ribosomal proteins. Overall, these results indicate that in the baseline transcriptome of future non-responders, many genes related to interferon and cytokine signaling are downregulated, whereas ribosomal proteins are upregulated.

### **Responders and non-responders have different expression levels of GPR15 and SEMA6B**

In order to identify the most robust genes predicting response to TNFi, we selected genes with the following features: mean RPKM>0.5 and an absolute fold change higher than 1.5 between responders and non-responders that is significant (nominal p-value <0.05). Overall 178 unique targets met these criteria in at least one of the six comparison considered (see list Supplementary Table 4). To identify the biological role(s) of these 178 identified genes, it was investigated whether they showed a relative overrepresentation of pathways (among those annotated in REACTOME, KEGG and Pubmed's GeneRIF) or GO-terms linked to biological processes. The enrichment analysis revealed that the 178 differentially expressed genes mainly belong to general immunologic pathways and biological processes (Table 2). For example, of all 1572 genes linked to the GO-term "immune response", 33 were present among the selection of 178 predictive genes (B&H adj. p-val = 7.06E-06). Yet also "Antigen binding" was overrepresented, with 10 genes being present in our selection out of 120 (B&H adj. p-val = 3.91E-06).

In order to identify specific targets that can be used for prediction of TNFi response, we identified those gene transcripts differentially expressed in multiple comparisons, out of the six considered (Supplementary Table 4). Overall we observed a small number of genes selected by the applied criteria: six genes were differentially expressed in all non-responders *versus* responders, and 22 genes in all extreme non-responders *versus* responders. The overlap in predictive genes between these two comparisons consisted of two targets: GPR15 and SEMA6B. Both these genes were also significantly different in ADA extreme non-responders *versus* extreme responders (among 102 genes in total), and SEMA6B also in all ADA non-responders *versus* all responders (among 31 genes in total). Three other genes (KLK1, MARC2 and RPL13P12) also overlapped over three comparisons, namely extreme non-responders *versus* responders, ADA non-responders *versus*

responders and ADA extreme non-responders *versus* responders. Since SEMA6B overlapped in four groups and GPR15 in the largest groups, we considered these the most interesting targets and focused further analyses on these two genes. The expression of GPR15 and SEMA6B as measured by RNA-seq is displayed in Figure 1. GPR15 showed a sensitivity of 55% and specificity of 75% to distinguish non-responders from responders, whereas SEMA6B showed a sensitivity of 37.5% and specificity of 92.5%, as determined by the Receiver Operating Characteristic (ROC) curve analysis (Supplementary Figure 1). These high specificities indicate that the implementation of SEMA6B and GPR15 measurement in a potential clinical test would be especially suitable to accurately identify non-responders, with the clinical implication of withholding treatment when the risk of response is extremely low.

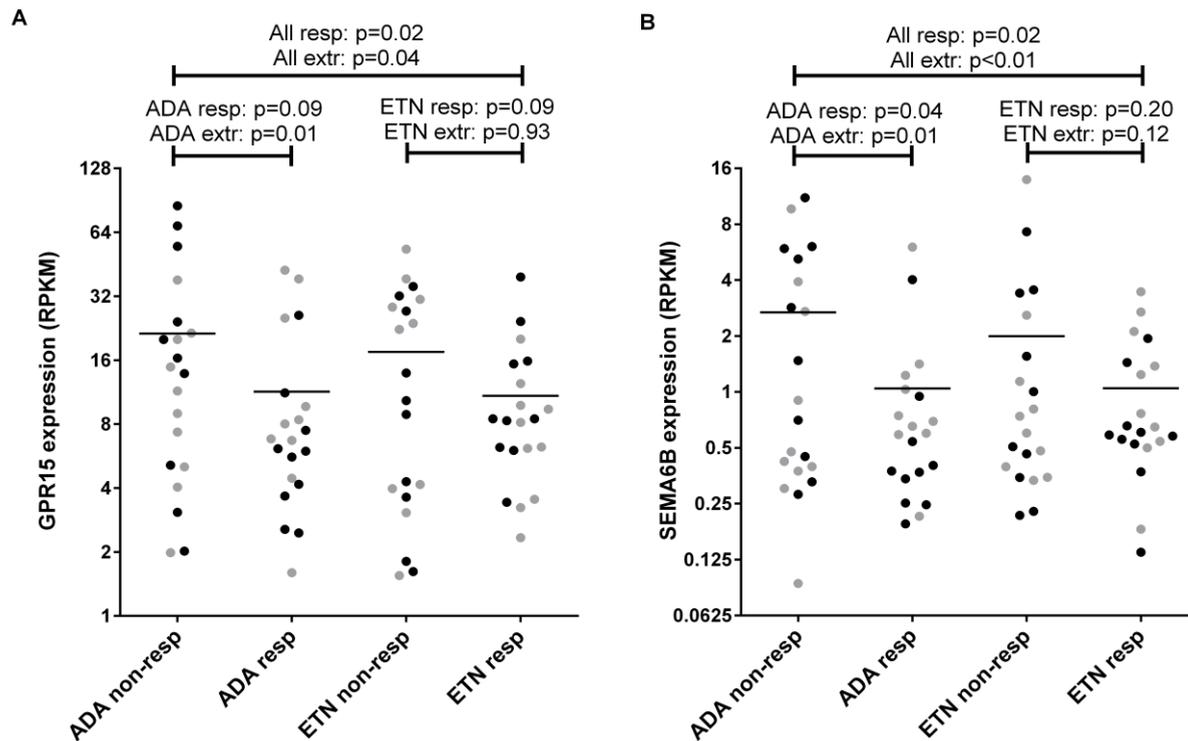
**Table 2:** Pathways and gene ontology (GO) biological processes related to the 178 selected genes. All 178 genes with a minimal expression (RPKM>0.5) and a significant difference between responders and non-responders ( $p<0.05$ ) with an absolute  $\log_2FC>0.58$ , were used as input for a Toppgene pathway enrichment analysis. 146/178 gene could be annotated to functional pathways, among which especially immune related pathways were overrepresented. Shown are the top 10 gene ontology (GO) molecular functions (MF)/biological processes (BP) and pathways, as based on nominal p-values (not shown). After a Benjamini & Hochberg correction of the p-value (B&H adj. P-val), all pathways in the top 10 were still significantly enriched.

Source	Pathway	B&H adjusted p-value	Genes total in pathway	No. of gene hits	Enriched genes
GeneRIF	New genetic associations detected in a host response study to hepatitis B vaccine.	2.05E-06	826	20	SLAMF8, COL4A3, FCRL2, HLA-B, HLA-C, FPR3, FCRLA, TNFRSF10C, ADORA3, ORM1, FCRL1, COCH, CD19, MS4A1, CD22, SIGLEC6, TNFRSF13C, C1QC, CD79A, CLEC10A
GeneRIF	Association study of B-cell marker gene polymorphisms in European Caucasian patients with systemic sclerosis.	2.05E-06	4	4	CD19, MS4A1, CD22, CD24
GO: BP	Immune response	7.06E-06	1572	33	MRC1, LILRB3, HLA-A, HLA-B, HLA-C, FOS, TNIP3, VSIG4, POU2AF1, TNFRSF10C, CD209, BLK, ADORA3, JUN, CD1E, THBS1, COCH, CD19, MS4A1, ERAP2, IGHD, IGHG1, IGHM, IGKC, CD24, TNFRSF13C, TNFRSF13B, C1QC, CCL3L1, CD79A, PAWR, PAX5, CLEC10A
GO: MF	Antigen binding	3.91E-06	120	10	HLA-A, HLA-B, HLA-C, CD209, CD1E, MS4A1, IGHD, IGHG1, IGHM, IGKC
GeneRIF	Association of maternal histocompatibility at class II HLA loci with maternal microchimerism in the fetus.	1.10E-05	7	4	HLA-A, HLA-B, HLA-C, GSTM1
GeneRIF	Functional characterization of the human immunodeficiency virus type 1 Nef acidic domain.	1.10E-05	10	4	HLA-A, HLA-B, HLA-C, JUN
GO: BP	Adaptive immune response	1.20E-04	402	15	LILRB3, HLA-A, HLA-B, CD209, CD1E, ERAP2, IGHD, IGHG1, IGHM, IGKC, TNFRSF13C, TNFRSF13B, C1QC, CD79A, CLEC10A
GO: BP	B cell receptor signaling pathway	1.98E-04	65	7	BLK, CD19, IGHD, IGHG1, IGHM, IGKC, CD79A

---

<b>GO: BP</b>	Antigen processing and presentation of endogenous peptide antigen via MHC class I via ER pathway, TAP-independent	1.98E-04	3	3	HLA-A, HLA-B, HLA-C
<b>GO: BP</b>	Regulation of immune response	3.27E-04	899	21	HLA-A, HLA-B, HLA-C, FOS, TNIP3, VSIG4, CD209, BLK, ADORA3, JUN, COCH, CD19, IGHD, IGHG1, IGHM, IGKC, CD24, TNFRSF13C, C1QC, CD79A, PAWR

---



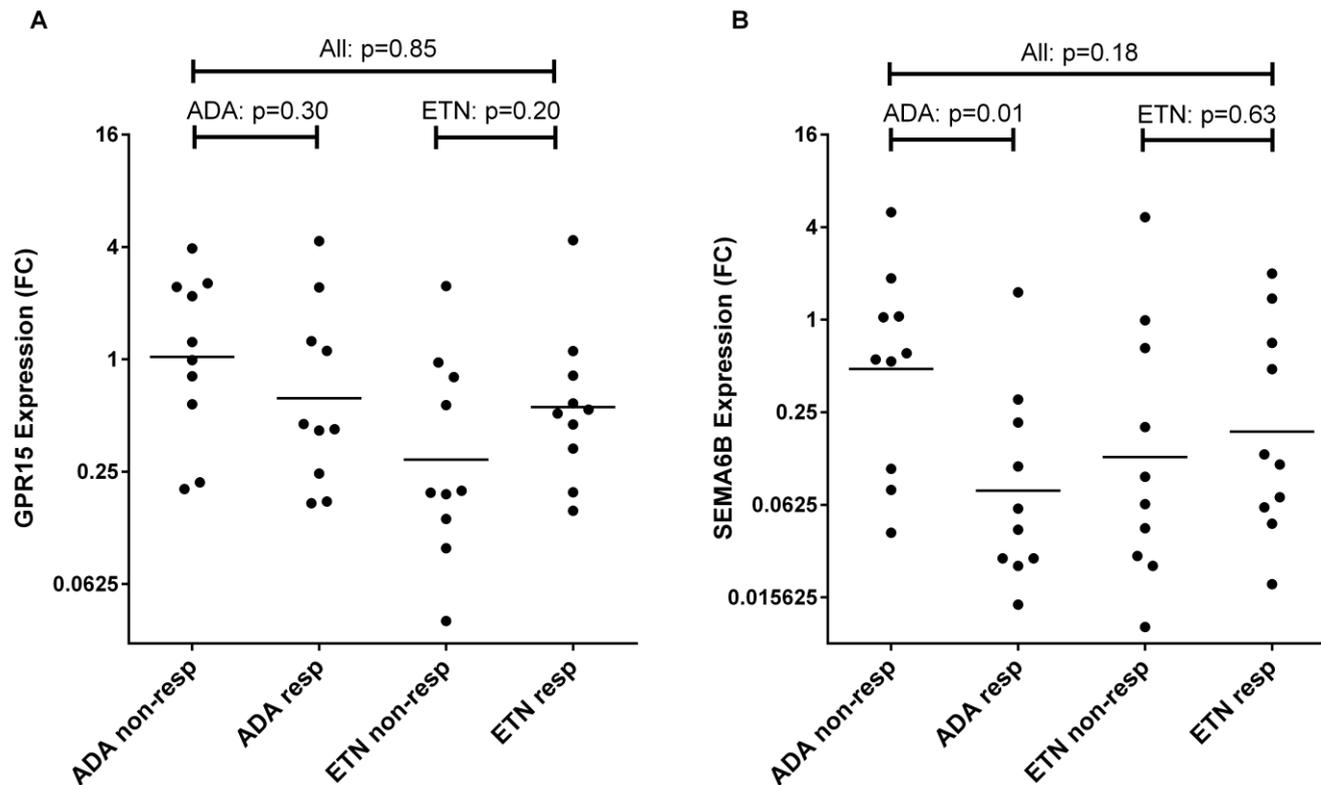
**Figure 1:** GPR15 and SEMA6B expression in the discovery cohort, split for ADA and ETN non-responders *versus* responders. Shown are the baseline gene expression levels of **(A)** GPR15 and **(B)** SEMA6B as measured by RNA-sequencing, in reads per kilobase per million mapped reads (RPKM), for non-responders and responders to treatment ( $n=40$  ADA and  $n=40$  ETN) in the discovery cohort. Black dots indicate the extreme (non-)responders among all patients, whereas the grey dots represent the remaining (non-)responders. The horizontal bar indicates the geometric mean.

### **GPR15 and SEMA6B are markers for smoking, rather than independent predictors**

In order to validate GPR15 and SEMA6B as predictors of response to TNFi therapy, the expression of GPR15 and SEMA6B was analyzed using quantitative PCR in the discovery cohort (i.e. technical replication, n=80) as well as in a new cohort of patients (validation, n=40). Technical replication in the discovery cohort showed comparable differences in expression for all comparisons as measured by qPCR instead of RNA-seq, yet in all cases the magnitude and statistical significance of the difference decreased to some extent (Table 3). Nevertheless, the observed contrast in the results was small enough to assume that this emerged from intrinsic differences between the two techniques used, which is also confirmed by the high correlation between the RNA-seq and qPCR measurements of 0.883 ( $p < 0.001$ ) and 0.857 ( $p < 0.001$ ) for GPR15 and SEMA6B respectively (Supplementary Figure 2). In the validation cohort, the expression levels of SEMA6B still showed a significant upregulation in ADA non-responders (FC 6.18,  $p = 0.01$ ), though in ETN an inverse direction as compared to the discovery phase was observed (FC=0.68,  $p = 0.63$ ) (Figure 2). GPR15 was also replicated as upregulated in non-responders to ADA, although not significant (FC=1.67,  $p = 0.30$ ), whereas in ETN also a downregulation was seen (FC=0.52,  $p = 0.20$ ).

**Table 3:** Technical replication of GPR15 and SEMA6B using single qPCR. In the discovery cohort of 80 patients, GPR15 and SEMA6B were selected as potential predictors for response to ADA and ETN. In order to test the technical replicability of these genes, all 80 samples were re-measured using single RT-qPCR assays. Shown are the fold changes (FC) of non-responders *versus* responders, and the corresponding p-value (based on DESeq-analysis for discovery, and t-test on  $-ddCt$  in the replication). The differential expression in the technical replication was in the same direction, however the FCs were in general less extreme, with a parallel increase in p-value.

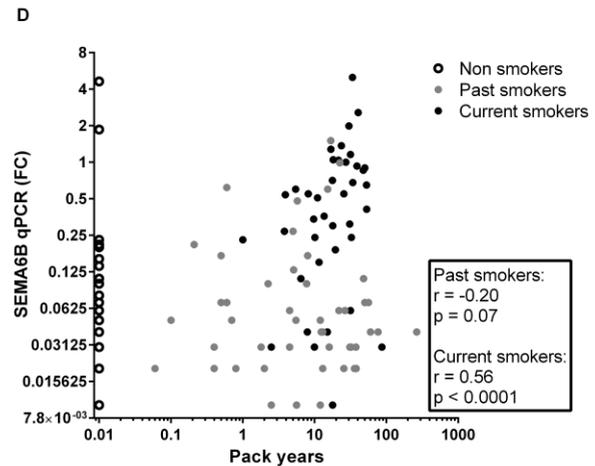
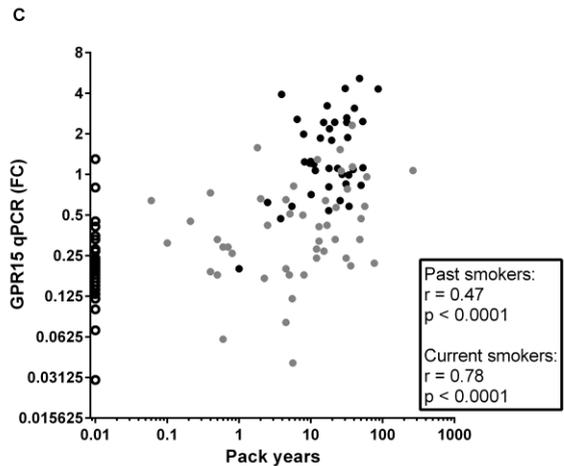
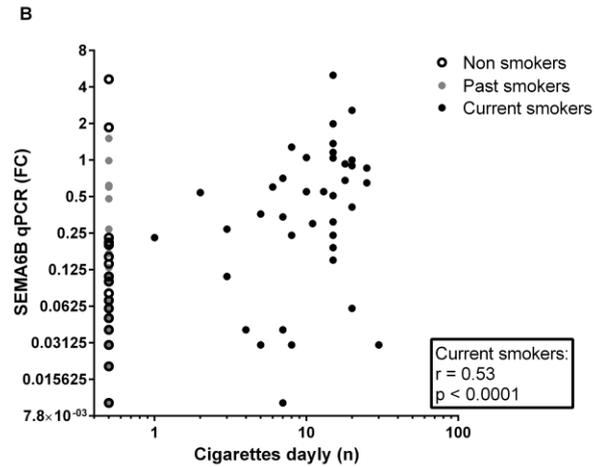
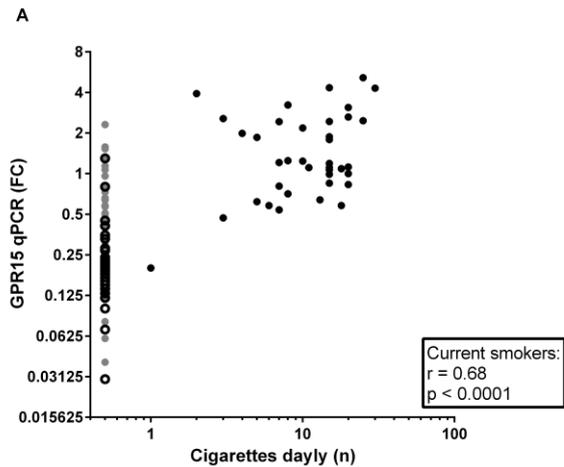
Subgroups	n	GPR15				SEMA6B			
		Discovery		Tech. replication		Discovery		Tech. Replication	
		FC	p-val	FC	p-val	FC	p-val	FC	p-val
All resp	80	1.752	<b>0.02</b>	1.494	0.08	2.190	<b>0.02</b>	1.432	0.27
All extr	20	2.073	<b>0.04</b>	1.514	0.27	3.603	<b>0.00</b>	2.723	<b>0.02</b>
Ada resp	40	1.873	0.09	1.559	0.18	2.431	<b>0.04</b>	1.476	0.41
Ada extr	20	3.891	<b>0.01</b>	2.977	<b>0.05</b>	4.516	<b>0.01</b>	4.238	<b>0.04</b>
Etn resp	40	1.628	0.09	1.432	0.28	1.941	0.20	1.390	0.37
Etn extr	20	1.086	0.93	0.770	0.59	2.681	0.12	1.750	0.30



**Figure 2:** GPR15 and SEMA6B expression in the validation cohort, split for ADA and ETN non-responders *versus* responders. For non-responders and responders to ADA ( $n=20$ ) and ETN ( $n=20$ ) treatment in the validation cohort, the baseline relative gene expression levels of (A) GPR15 and (B) SEMA6B as measured by RT-qPCR, is shown as the fold change (FC). The horizontal bar indicates the geometric mean.

The ability to validate SEMA6B and (to some extent) GPR15 as predictors of ADA response, though not of ETN response, raised the question whether other parameters were of influence in the relation between the expression levels of these transcripts and the response to therapy. Therefore we additionally investigated if the expression of SEMA6B and GPR15 could be a reflection of the PBMC cell-subset composition before mRNA extraction. For the 120 samples in the discovery and validation cohort (full cohort), the correlation between gene expression and the percentage of cell subsets as measured by FACS was calculated (Supplementary Figures 3 & 4). SEMA6B expression did not significantly correlate with any cell-subset, whereas GPR15 expression very weakly correlated with the relative number of monocytes ( $r = -0.20$ ,  $p=0.04$ ) and B-cells ( $r = 0.25$ ,  $p=0.01$ ), indicating that PBMC composition probably had no influence on the results. Additionally, we verified whether the expression of these genes was associated with any baseline clinical characteristics of the patients enrolled in the study. Higher expression SEMA6B was most notably associated with current smoking ( $B=2.41$ ,  $p=0.000$ ) and the CRP levels ( $B=0.32$   $p=0.029$ ), whereas higher expression of GPR15 was associated with current smoking ( $B=2.18$ ,  $p=0.000$ ) and lower expression with concomitant glucocorticoid (GC)-use ( $B=-0.551$ ,  $p=0.015$ ). The association between smoking and the higher expression of GPR15 and SEMA6B explained why these genes were identified as robustly differentially expressed only in the discovery phase, as in the discovery cohort current smokers are 15% more represented among non-responders to ADA and ETN (Supplementary Table 1). In line with this observation, the ability to validate SEMA6B and (to some extent) GPR15 in ADA can be explained by a 30% higher proportion of current smokers among non-responders, whereas the inability to validate these genes for ETN can be explained by the relative underrepresentation of smokers (20% less) among ETN non-responders in the validation cohort. We further investigated the relationship between the expression of these two genes and smoking, by stratifying the patients into never smokers, past smokers and current smokers, while concomitantly accounting for the number of daily smoked cigarettes and total smoked pack years (PY) (Figure 3). While focusing on current smokers, we observed positive correlations with extremely significant p-values ( $p<0.0001$ ) between the number of daily smoked cigarettes and PY with the expression of GPR15 and SEMA6B, thus indicating that there is a dose-response relationship for the exposure to smoke and

upregulation of these genes. Additionally, the highly significant positive correlation between PY and expression of GPR15 in past smokers, indicates that there is a certain prolonged and cumulative effect of (heavy) smoking and the upregulation of this gene, whereas for SEMA6B these effects were not observed.



**Figure 3:** Correlations between smoking and GPR15 and SEMA6B expression. Graphs show the correlation between the relative gene expression levels of GPR15 and SEMA6B as measured by single qPCR assay and (A & B) the number of cigarettes smoked per day and (C & D) pack years (PY) of smoking (i.e. 1 PY = 20 cigarettes per day for one year). Patients were categorized as current smokers, discontinued smokers and never-smokers, the latter of which were given arbitrarily the lowest value for daily cigarettes or pack years. All FC values >1 were in (discontinued) smokers, except for three patients who declared to have never smoked. These three patients were contacted in retrospect to ask if there was any possibility of smoke inhalation at the time of the baseline visit. Two patients declared that they did not inhale smoke around treatment initiation, whereas one (with GPR15 FC = 0.8 and SEMA6B FC = 4.6) declared to be occasionally in a room with a smoker. A clear correlation for daily cigarettes and pack years with GPR15 and SEMA6B expression was seen in current smokers, and for pack years and GPR15 expression in discontinued smokers (all  $p < 0.0001$ ).

## DISCUSSION

In this study, we investigated the baseline transcriptome of PBMC mRNA for predictors of response to TNFi. GPR15 and SEMA6B were revealed as the most promising gene transcripts for prediction of response to ADA and ETN, of which SEMA6B could be validated in a separate cohort of consecutively included ADA treated patients. However, additional analyses revealed that the expression of both genes was strongly associated with smoking, which was unequally distributed between responders and non-responders. It has been frequently reported, although not consistently, that smoking decreases the chances on response, remission and continuation of treatment in RA [29-34]. The exact mechanism behind this observation remains unclear, however, it is possible that SEMA6B and GPR15 play a role in the mechanism that leads to a reduced therapeutic effect of TNFi. This hypothesis is not unlikely, given that SEMA6B and GPR15 are strongly and dose-dependently related to smoking and probably play a relevant role in the pathophysiology of RA.

SEMA6B is a protein with pro-proliferative effects via its receptor Plexin-A4, which forms complexes with fibroblast growth factor-receptor 1 and 2 (FGFR1-2) and vascular endothelial growth factor-receptor 2 (VEGFR2), thereby inducing the signal transduction of these receptors [35]. VEGF and bFGF are the ligands of these receptors, and are known to have instrumental roles in RA: VEGF is a critical angiogenesis factor responsible for vascular proliferation and blood vessel invasion of the synovial lining membrane in RA, whereas acidic FGF (FGF-1) and basic FGF (FGF-2) have also been implicated in synovial hyperplasia and apoptosis resistance in adult RA [36]. Considering that SEMA6B has the ability to influence these important proinflammatory pathways in RA pathogenesis, its upregulation may explain the insensitivity to treatment. Additionally and in line with this study, upregulation of SEMA6B in smokers has been described before in a genome-wide study on PBMC mRNA ( $p < 0.0001$ ) [37], while no association was found in whole blood mRNA [38], and no relation between smoking and DNA methylation of the SEMA6B gene has been reported [39-42].

In turn, GPR15 or Brother of Bonzo (BOB) is expressed on the cell surface of monocytes and neutrophils, and is a chemo attractant for T-cells [43, 44]. GPR15 mRNA expression is increased in the synovium and in the peripheral blood leukocytes of RA patients compared to non-RA controls [44, 45]; similarly, the protein expression of GPR15 is increased in synovial tissue macrophages and in circulating monocytes and neutrophils [44]. In one of the previous microarray studies investigating prediction of response [46], GPR15 was clustered within the “M1 activated monocytes” group that was more strongly associated with good versus poor response to TNFi therapy than other clusters (p-value=0.006). The dysregulation of GPR15 in different cell subsets of RA patients and its association with TNFi response in the current and previous study, suggest that GPR15 might be a biologically relevant predictor. Additionally, we found a clear dose-response effect of smoking on the expression of GPR15, which also showed a cumulative and prolonged effect, as the expression of GPR15 was still upregulated in past smokers. The strong correlation between smoking and a higher expression of GPR15 is also supported by literature, as smokers repeatedly show DNA hypomethylation of the GPR15 gene and subsequent increased mRNA expression [39-42, 47, 48]. Because of its biological role and the strong relation with smoking, GPR15 is a good candidate to explain the health hazards of smoking with regard to chronic inflammatory disease [48], and in particular the reduced therapeutic effect to TNFi.

Altogether, given our results and current knowledge from literature, we can speculate that smoking dose dependently increases GPR15 and SEMA6B expression in PBMCs and thereby directly or indirectly induces respectively T-cell involvement and synovial vascular proliferation, leading to a more severe and treatment refractory RA. However, as the exact upstream and downstream pathways related to these genes are largely unknown, the possibility to investigate if the aberrant expression of SEMA6B and GPR15 reflects a larger dysregulation on the entire transcriptome using these data is limited. Future functional *in-vivo* studies should address whether the mRNA and protein expression of SEMA6B and GPR15 are modulated in the synovial tissue (ST) of a joint before and after (heavy) smoke exposure. The specific link between SEMA6B and GPR15 on T-cell

recruitment and synovial vascular proliferation, and subsequent RA severity and non-response to treatments, represent another but more challenging topic to investigate.

Interestingly, neither the TNF $\alpha$  nor LT $\alpha$ -pathway self, but rather general immunologic pathways at baseline were found to be related to a decreased response to ADA and ETN. This finding supports the hypothesis, already suggested by [49], that TNFi's (and non-TNFi's) eventually mediate their efficacy by interfering with a common final pathway, namely proinflammatory cytokine production. Among the most altered pathways, we identified three interferon-related pathways significantly downregulated in non-responders at baseline. In line with this observation, are the results Wright and colleagues [14], who compared the expression of genes in neutrophils of RA patients (before start of treatment) and non-RA control patients by RNA-seq, and found the IFN-signaling pathway as most indicative for RA. Subsequently, upregulation of an IFN-score (calculated on the expression levels of these and other IFN-related genes) predicted EULAR good response to TNFi (ADA, ETN and golimumab) in these RA patients, with an area under the receiver operating characteristic curve (AUC-ROC) of 0.76. Although these results were not attempted to be replicated in a separate cohort, in our selection of 178 genes predictive of response, five genes overlapped with those included in the IFN-score: CCL3L3, THBS1, HLA-A, HLA-B and HLA-C. However, the proposed IFN-score by Wright et al. calculated on the basis of our PBMC RNA-seq data-set did not predict the TNFi response (AUC=0.55, 95% CI=0.42-0.68) (Supplementary Figure 5). The inconsistencies between the results of our study and those produced by and Wright et al. could be related to the starting cell source (PBMC *versus* neutrophils), considering that neutrophils have been identified as biggest contributors in type I IFN signature seen in RA [50]. In addition, it should also be pointed out that Wright et al. did not cross-validated their findings in another independent cohort of patients, of which the importance has previously been reported for this type of studies [51]. The design was therefore a strong point of the study presented here, where a biomarker discovery was initially performed in a large cohort of patients with a clear distinctive clinical outcome, and the reliability of findings was tested by performing a technical replication and a biological validation in a distinct cohort of patients.

## **CONCLUSION**

Efforts of the present and previous studies have not provided gene transcripts that are able to independently and repeatedly predict response to TNFi treatment. However, we demonstrated that pathways altered in non-responders at baseline are rather linked to general immune functions and cytokines than to TNF $\alpha$  and LT $\alpha$  specifically, indicating that all biological treatments target a common cytokine related pathway, as previously suggested by others. Furthermore, the identification of GPR15 and SEMA6B expression as markers of response and (substitute) dose-dependent indicators of smoking, opens new venues for the identification of the molecular mechanisms underlying TNFi refractory RA.

## REFERENCES

1. Cuppen BV, Welsing PM, Sprengers JJ, Bijlsma JW, Marijnissen AC, van Laar JM et al. Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability. *Rheumatology (Oxford)* 2016;55 5: 826-39.
2. Smith SL, Plant D, Eyre S, Barton A. The potential use of expression profiling: implications for predicting treatment response in rheumatoid arthritis. *Ann Rheum Dis* 2013;72 7: 1118-24.
3. Burska AN, Roget K, Blits M, Soto Gomez L, van de Loo F, Hazelwood LD et al. Gene expression analysis in RA: towards personalized medicine. *Pharmacogenomics J* 2014;14 2: 93-106.
4. Toonen EJ, Gilissen C, Franke B, Kievit W, Eijssbouts AM, den Broeder AA et al. Validation study of existing gene expression signatures for anti-TNF treatment in patients with rheumatoid arthritis. *PLoS One* 2012;7 3: e33199.
5. MacIsaac KD, Baumgartner R, Kang J, Loboda A, Peterfy C, DiCarlo J et al. Pre-treatment whole blood gene expression is associated with 14-week response assessed by dynamic contrast enhanced magnetic resonance imaging in infliximab-treated rheumatoid arthritis patients. *PLoS One* 2014;9 12: e113937.
6. Thomson TM, Lescarbeau RM, Drubin DA, Laifenfeld D, de GD, Fryburg DA et al. Blood-based identification of non-responders to anti-TNF therapy in rheumatoid arthritis. *BMC Med Genomics* 2015;8: 26.
7. Malone JH, Oliver B. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol* 2011;9: 34.
8. Sirbu A, Kerr G, Crane M, Ruskin HJ. RNA-Seq vs dual- and single-channel microarray data: sensitivity analysis for differential expression and clustering. *PLoS One* 2012;7 12: e50986.
9. Hurd PJ, Nelson CJ. Advantages of next-generation sequencing versus the microarray in epigenetic research. *Brief Funct Genomic Proteomic* 2009;8 3: 174-83.
10. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 2008;5 7: 621-8.
11. Fu X, Fu N, Guo S, Yan Z, Xu Y, Hu H et al. Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics* 2009;10: 161.
12. 't Hoen PA, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RH, de Menezes RX et al. Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Res* 2008;36 21: e141.
13. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* 2008;18 9: 1509-17.
14. Wright HL, Thomas HB, Moots RJ, Edwards SW. Interferon gene expression signature in rheumatoid arthritis neutrophils correlates with a good response to TNFi therapy. *Rheumatology (Oxford)* 2015;54 1: 188-93.
15. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38 1: 44-8.
16. van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996;39 1.
17. Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K et al. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 2009;25 15: 1966-7.
18. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;31 2: 166-9.
19. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010;11 10: R106.
20. Mayeux R. Biomarkers: potential uses and limitations. *NeuroRx* 2004;1 2: 182-8.
21. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta Delta C(T)) Method. *Methods* 2001;25 4: 402-8.
22. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 2016;44 D1: D457-62.
23. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al. Gene set enrichment analysis: a knowledge-based

- approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102 43: 15545-50.
24. Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R et al. The Reactome pathway Knowledgebase. *Nucleic Acids Res* 2016;44 D1: D481-7.
  25. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 2009;37 Web Server issue: W305-11.
  26. Youden W, . Index for rating diagnostic tests. *Cancer* 1950;3 1: 32-5.
  27. Tansey MG, Szymkowski DE. The TNF superfamily in 2009: new pathways, new indications, and new drugs. *Drug Discov Today* 2009;14 23-24: 1082-8.
  28. Tracey D, Klareskog L, Sasso EH, Salfeld JG, Tak PP. Tumor necrosis factor antagonist mechanisms of action: a comprehensive review. *Pharmacol Ther* 2008;117 2: 244-79.
  29. Abhishek A, Butt S, Gadsby K, Zhang W, Deighton CM. Anti-TNF-(alpha) agents are less effective for the treatment of rheumatoid arthritis in current smokers. *J Clin Rheumat* 2010;16 1: 15-8.
  30. Forslind K, Hafstrom I, Ahlmen M, Svensson B, Group BS. Sex: a major predictor of remission in early rheumatoid arthritis? *Ann Rheum Dis* 2007;66 1: 46-52.
  31. Hyrich KL, Watson KD, Silman AJ, Symmons DP. 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: 1558-65.
  32. Saevarsdottir S, Wedren S, Seddighzadeh M, Bengtsson C, Wesley A, Lindblad S et al. Patients with early rheumatoid arthritis who smoke are less likely to respond to treatment with methotrexate and tumor necrosis factor inhibitors: observations from the Epidemiological Investigation of Rheumatoid Arthritis and the Swedish Rheumatology Register cohorts. *Arthritis Rheum* 2011;63 1: 26-36.
  33. Matthey DL, Brownfield A, Dawes PT. Relationship between pack-year history of smoking and response to tumor necrosis factor antagonists in patients with rheumatoid arthritis. *J Rheumatol*. 2009;1180-7.
  34. Cuppen BV, Jacobs JW, Ter Borg EJ, Marijnissen AC, Bijlsma JW, Lafeber FP et al. Necessity of TNF-alpha inhibitor discontinuation in rheumatoid arthritis is predicted by smoking and number of previously used biological DMARDs. *Clin Exp Rheumatol* 2016.
  35. Kigel B, Rabinowicz N, Varshavsky A, Kessler O, Neufeld G. Plexin-A4 promotes tumor progression and tumor angiogenesis by enhancement of VEGF and bFGF signaling. *Blood* 2011;118 15: 4285-96.
  36. Malemud CJ. Growth hormone, VEGF and FGF: involvement in rheumatoid arthritis. *Clin Chim Acta* 2007;375 1-2: 10-9.
  37. Charlesworth JC, Curran JE, Johnson MP, Goring HH, Dyer TD, Diego VP et al. Transcriptomic epidemiology of smoking: the effect of smoking on gene expression in lymphocytes. *BMC Med Genomics* 2010;3: 29.
  38. Beineke P, Fitch K, Tao H, Elashoff MR, Rosenberg S, Kraus WE et al. A whole blood gene expression-based signature for smoking status. *BMC Med Genomics* 2012;5: 58.
  39. Tsaprouni LG, Yang TP, Bell J, Dick KJ, Kanoni S, Nisbet J et al. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics* 2014;9 10: 1382-96.
  40. Wan ES, Qiu W, Baccarelli A, Carey VJ, Bacherman H, Rennard SI et al. Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. *Hum Mol Genet* 2012;21 13: 3073-82.
  41. Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet* 2011;88 4: 450-7.
  42. Gao X, Jia M, Zhang Y, Breitling LP, Brenner H. DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. *Clin Epigenetics* 2015;7: 113.
  43. Kim SV, Xiang WV, Kwak C, Yang Y, Lin XW, Ota M et al. GPR15-mediated homing controls immune homeostasis in the large intestine mucosa. *Science* 2013;340 6139: 1456-9.
  44. Cartwright A, Schmutz C, Askari A, Kuiper JH, Middleton J. Orphan receptor GPR15/BOB is up-regulated in rheumatoid arthritis. *Cytokine* 2014;67 2: 53-9.
  45. Schmutz C, Hulme A, Burman A, Salmon M, Ashton B, Buckley C et al. Chemokine receptors in the rheumatoid synovium: upregulation of CXCR5. *Arthritis Res Ther* 2005;7 2: R217-29.
  46. Dennis G, Jr., Holweg CT, Kummerfeld SK, Choy DF, Setiadi AF, Hackney JA et al. Synovial phenotypes in rheumatoid arthritis correlate

- with response to biologic therapeutics. *Arthritis Res Ther* 2014;16 2: R90.
47. Dogan MV, Xiang J, Beach SR, Cutrona C, Gibbons FX, Simons RL et al. Ethnicity and Smoking-Associated DNA Methylation Changes at HIV Co-Receptor GPR15. *Front Psychiatry* 2015;6: 132.
  48. Koks G, Uudelepp ML, Limbach M, Peterson P, Reimann E, Koks S. Smoking-induced expression of the GPR15 gene indicates its potential role in chronic inflammatory pathologies. *Am J Pathol* 2015;185 11: 2898-906.
  49. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet* 2016.
  50. de Jong TD, Lubbers J, Turk S, Vosslamber S, Mantel E, Bontkes HJ et al. The type I interferon signature in leukocyte subsets from peripheral blood of patients with early arthritis: a major contribution by granulocytes. *Arthritis Res Ther* 2016;18: 165.
  51. Cuppen BV, Rossato M, Fritsch-Stork RD, Concepcion AN, Schenk Y, Bijlsma JW et al. Can baseline serum microRNAs predict response to TNF-alpha inhibitors in rheumatoid arthritis? *Arthritis Res Ther* 2016;18: 189.

**Supplementary Table 1:** Baseline characteristics of patients in discovery (n=80) and validation (n=40) cohort, split for treatment and response. Patients were selected from the observational BiOCURA cohort based on treatment outcome over the course of one year after start of either ADA or ETN. The presented clinical characteristics for responders and non-responders are all before treatment initiation. P-values for comparisons between responders and non-responders were calculated by means of an independent sample t-test, Mann-Whitney U test, fisher exact test (2\*2) or chi-square (>2\*2) based on distribution of the clinical parameter.

### 1a Discovery cohort

Item	ADA (n=40)			ETN (n=40)		
	Non-resp (n=20)	Resp (n=20)	p-value	Non-resp (n=20)	Resp (n=20)	p-value
Female gender, n (%)	15 (75)	14 (70)	1.00	17 (85)	15 (75)	0.70
Age, mean years $\pm$ sd	53.4 $\pm$ 11.9	56.4 $\pm$ 12.9	0.46	56.3 $\pm$ 9.7	53.9 $\pm$ 10.1	0.44
Current smoker, n (%)	8 (40)	5 (25)	0.50	7 (35)	4 (20)	0.48
RF positivity, n (%)	11 (55)	16 (80)	0.18	14 (70)	15 (75)	0.72
ACPA positivity, n (%)	12 (60)	15 (75)	0.50	14 (70)	17 (85)	0.13
CRP, median	10.0 (4.0-13.0)	7.5 (4.0-12.8)	0.97	4.3 (3.1-7.5)	8.5 (3.8-30.0)	0.07
No. of prev. bDMARDs			0.14			0.57
0, n (%)	13 (65)	17 (85)		15 (75)	15 (75)	
1, n (%)	7 (35)	3 (15)		4 (20)	5 (25)	
2, n (%)	0 (0)	0 (0)		1 (5)	0 (0)	
Concomitant treatment, n (%)	19 (95)	19 (95)	1.00	20 (100)	19 (95)	1.00
MTX, n (%)	13 (65)	17 (85)	0.27	13 (65)	16 (80)	0.48
SSZ, n (%)	2 (10)	3 (15)	1.00	3 (15)	1 (5)	0.61
HCQ (%)	5 (25)	5 (25)	1.00	7 (35)	6 (30)	1.00
GCs, n (%)	8 (40)	3 (15)	0.16	7 (35)	5 (25)	0.73
Baseline DAS28, mean $\pm$ sd	4.3 $\pm$ 1.3	4.8 $\pm$ 1.0	0.15	4.4 $\pm$ 1.2	4.8 $\pm$ 1.0	0.28
TJC, median (IQR)	7.0 (3.0-13.8)	7.0 (4.3-14.5)	0.88	4.5 (1.0-21.5)	6.5 (2.0-17.8)	1.00
SJC, median (IQR)	1.0 (0.0-2.8)	3.0 (0.0-6.8)	0.08	1.0 (0.0-3.0)	1.5 (0.3-4.8)	0.30
VAS, mean $\pm$ sd	62.0 $\pm$ 20.7	64.5 $\pm$ 23.8	0.73	54.5 $\pm$ 26.0	54.5 $\pm$ 23.9	1.00
ESR, median mm/hr (IQR)	12.5 (2.5-30.0)	18.5 (9.0-29.8)	0.33	20.0 (5.3-35.3)	27.0 (13.0-54.3)	0.23

### 1b Validation cohort

Item	ADA (n=20)			ETN (n=20)		
	Non-resp (n=10)	Resp (n=10)	p-value	Non-resp (n=10)	Resp (n=10)	p-value
Female gender, n (%)	6 (60)	7 (70)	1.00	8 (80)	6 (60)	0.63
Age, mean years $\pm$ sd	56.4 $\pm$ 8.9	47.6 $\pm$ 10.8	0.06	62.3 $\pm$ 6.8	57.7 $\pm$ 11.5	0.28
Current smoker, n (%)	7 (70)	4 (40)	0.37	1 (10)	3 (30)	0.58
RF positivity, n (%)	5 (50)	5 (50)	1.00	6 (60)	6 (60)	1.00
ACPA positivity, n (%)	7 (70)	4 (40)	0.37	5 (50)	8 (80)	0.35
CRP, median	1.6 (1.0-6.8)	2.7 (1.0-10.8)	0.80	2.0 (1.0-23.5)	8.0 (3.3-12.3)	0.53
No. of prev. bDMARDs			0.42			0.55
0, n (%)	7 (70)	6 (60)		7 (70)	7 (70)	
1, n (%)	2 (20)	4 (40)		3 (30)	2 (20)	
2, n (%)	1 (10)	0 (0)		0 (0)	1 (10)	
Concomitant treatment, n (%)	10 (100)	10 (100)	-	7 (70)	10 (100)	0.21
MTX, n (%)	8 (80)	10 (100)	0.47	5 (50)	9 (90)	0.14
SSZ, n (%)	0 (0)	1 (10)	1.00	1 (10)	1 (10)	1.00
HCQ (%)	3 (30)	2 (20)	1.00	3 (30)	5 (50)	0.65
GCs, n (%)	6 (60)	1 (10)	0.06	4 (40)	1 (10)	0.30
Baseline DAS28, mean $\pm$ sd	3.2 $\pm$ 1.3	4.5 $\pm$ 0.7	<b>0.02</b>	4.1 $\pm$ 1.2	4.5 $\pm$ 0.8	0.55
TJC, median (IQR)	2.0 (0.0-11.5)	5.5 (3.5-14.3)	0.12	7.0 (2.8-15.0)	4.0 (3.0-10.0)	0.63
SJC, median (IQR)	0.0 (0.0-0.0)	0.5 (0.0-3.0)	0.12	1.0 (0.0-4.0)	2.0 (0.8-3.0)	0.48
VAS, mean $\pm$ sd	41.5 $\pm$ 24.7	62.5 $\pm$ 18.7	0.05	57.5 $\pm$ 15.9	52.0 $\pm$ 23.0	0.54
ESR, median mm/hr (IQR)	9.0 (4.5-21.3)	14.0 (9.3-38.8)	0.25	9.5 (5.3-22.5)	18.5 (13.0-31.5)	0.17

ACPA, anti-citrullinated protein antibody; ADA, adalimumab; bDMARDs, biological disease modifying antirheumatic drugs; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; ETN, etanercept; GC, glucocorticoid; HCQ, hydroxychloroquine; IQR, interquartile range; MTX, methotrexate; RF, rheumatoid factor; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count; VAS-GH, visual analogue scale general health.

**Supplementary Table 2:** Differential expression of genes in pathways activated downstream the TNF-RI, TNF-RII and LT- $\beta$ R. In the discovery cohort of 80 patients, gene transcripts were measured by RNA-seq and tested for differential expression between non-responders and responders to ADA and ETN treatment. Known targets downstream of TNF-RI, TNF-RII and LT- $\beta$ R were derived from KEGG (ID: 04668 and 04064) and these gene transcripts were extracted from these results. The table reports the relative expression of each gene as relative fold change (FC) in non-responders *versus* responders and the corresponding B&H adj. p-value as calculated by DESeq. Significant results with a p-val < 0.05 are shown in bold.

Pathway	Gene	All resp		All extr		ADA resp		ADA extr		ETN resp		ETN extr	
		FC	p-val	FC	p-val	FC	p-val	FC	p-val	FC	p-val	FC	p-val
TNF-RI	AKT1	1.001	0.968	0.987	0.888	0.995	0.980	0.976	0.817	1.007	0.933	0.999	0.990
/TNF-RII	AKT2	1.009	0.867	0.988	0.913	1.022	0.801	1.023	0.858	0.996	0.993	0.954	0.750
	AKT3	1.022	0.761	1.007	0.901	1.034	0.728	1.125	0.358	1.009	0.932	0.899	0.483
	ATF2	1.008	0.948	1.027	0.803	1.014	0.909	1.011	0.924	1.002	0.981	1.044	0.779
	ATF4	1.064	0.372	1.031	0.733	1.049	0.617	1.088	0.557	1.079	0.406	0.977	0.922
	ATF6B	1.778	0.146	2.732	0.135	1.890	0.231	2.297	0.391	1.613	0.440	3.608	0.199
	ATF6B	1.255	0.721	1.285	0.544	1.118	0.844	1.535	0.471	2.107	0.690	1.076	0.936
	ATF6B	1.106	0.759	1.571	0.733	0.988	1.000	1.025	1.000	1.097	0.824	2.664	0.684
	BAG4	0.992	0.826	0.980	0.792	1.019	0.914	0.990	0.904	0.965	0.654	0.972	0.797
	BCL3	0.908	0.458	0.899	0.476	0.915	0.670	1.001	0.966	0.902	0.526	0.804	0.278
	BIRC2	0.957	0.528	0.982	0.822	0.963	0.687	0.940	0.615	0.951	0.587	1.026	0.881
	BIRC3	1.073	0.348	1.107	0.298	1.128	0.241	1.328	<b>0.021</b>	1.021	0.839	0.918	0.521
	CASP10	0.958	0.548	0.978	0.781	1.054	0.603	1.013	0.938	0.869	0.143	0.947	0.641
	CASP3	0.967	0.571	1.066	0.584	1.035	0.763	1.070	0.659	0.897	0.238	1.064	0.760
	CASP7	0.962	0.580	1.010	0.987	0.936	0.487	0.884	0.310	0.989	0.906	1.151	0.340
	CASP8	1.063	0.387	1.056	0.571	1.057	0.555	1.100	0.438	1.069	0.486	1.013	0.926
	CCL2	0.435	0.325	0.538	0.096	1.285	0.769	0.474	0.145	0.183	0.183	0.623	0.430
	CCL20	0.817	0.455	0.935	0.812	0.622	0.288	0.875	0.765	1.112	0.780	1.012	1.000
	CCL5	0.965	0.884	0.884	0.571	0.900	0.641	0.740	0.307	1.032	0.821	1.036	0.860
	CEBPB	0.902	0.367	0.915	0.549	0.914	0.632	0.952	0.773	0.890	0.403	0.882	0.608
	CFLAR	0.979	0.774	0.982	0.856	0.971	0.761	0.973	0.812	0.987	0.904	0.991	0.970
	CHUK	0.956	0.517	1.000	0.967	0.964	0.716	0.958	0.732	0.948	0.545	1.044	0.796
	CREB1	1.008	0.937	1.014	0.885	1.027	0.816	1.036	0.778	0.989	0.900	0.993	0.958
	CREB3	0.997	0.982	0.989	0.922	1.004	0.958	0.935	0.604	0.990	0.927	1.044	0.743
	CREB3L1	0.893	0.574	0.881	0.657	1.007	1.000	1.212	0.815	0.793	0.455	0.681	0.359
	CREB3L2	1.022	0.751	1.031	0.772	0.989	0.944	0.989	0.920	1.057	0.578	1.077	0.595
	CREB3L3	1.117	0.822	0.738	0.600	0.714	0.612	0.498	0.496	1.659	0.378	1.090	1.000

CREB3L4	1.046	0.532	1.084	0.430	1.106	0.320	1.153	0.284	0.990	0.929	1.021	0.893
CREB5	0.963	0.603	1.087	0.587	0.990	0.881	1.060	0.774	0.937	0.566	1.114	0.648
CSF1	0.941	0.434	0.956	0.718	0.940	0.512	1.047	0.690	0.942	0.627	0.879	0.474
CSF2	2.826	<b>0.036</b>	2.237	0.321	1.963	0.428	1.724	0.797	4.216	0.067	2.958	0.423
CX3CL1	1.000	0.985	0.813	0.283	0.930	0.763	0.860	0.581	1.074	0.708	0.769	0.387
CXCL1	0.739	0.513	0.633	0.341	0.357	0.198	0.630	0.458	1.436	0.412	0.635	0.538
CXCL10	0.457	0.564	1.740	0.219	3.230	0.245	0.845	0.433	0.142	0.308	3.304	0.106
CXCL2	1.008	0.925	1.219	0.708	0.624	0.622	1.465	0.666	2.114	0.075	0.925	0.961
CXCL3	0.828	0.803	1.047	0.938	0.553	0.570	1.344	0.720	1.499	0.245	0.744	0.636
CXCL5	1.221	0.398	1.350	0.156	1.446	0.369	1.232	0.313	1.037	0.887	1.470	0.296
DNM1L	0.995	0.924	0.988	0.894	0.993	0.921	0.977	0.844	0.996	0.965	1.000	0.998
EDN1	0.998	0.909	0.982	1.000	0.951	0.857	1.242	0.282	1.046	0.694	0.762	0.401
FADD	0.905	0.198	0.992	0.871	0.929	0.488	0.929	0.639	0.880	0.268	1.060	0.806
FAS	0.962	0.579	0.910	0.287	0.988	0.885	0.902	0.390	0.937	0.492	0.919	0.462
FOS	1.349	0.087	1.113	0.660	1.016	0.931	0.859	0.579	1.795	<b>0.019</b>	1.463	0.298
ICAM1	0.953	0.608	1.000	0.985	0.973	0.815	1.114	0.657	0.932	0.610	0.887	0.646
IKKBK	0.972	0.687	0.985	0.857	0.990	0.922	0.984	0.884	0.953	0.605	0.987	0.904
IKBKG	0.979	0.776	0.974	0.764	1.017	0.856	0.981	0.863	0.941	0.526	0.967	0.789
IL15	0.981	0.764	1.104	0.372	1.103	0.456	1.082	0.570	0.869	0.306	1.128	0.467
IL18R1	0.889	0.123	0.977	0.817	0.785	<b>0.018</b>	0.888	0.402	1.013	0.839	1.082	0.517
IL1B	0.765	0.236	0.631	0.077	0.667	0.250	0.772	0.391	0.865	0.638	0.496	0.113
IL6	1.262	0.378	0.979	0.997	0.928	0.845	1.433	0.303	1.546	0.273	0.755	0.714
ITCH	1.003	0.965	1.044	0.657	1.013	0.890	1.024	0.834	0.994	0.938	1.065	0.652
JAG1	0.993	0.917	1.090	0.488	1.037	0.773	1.134	0.406	0.948	0.525	1.050	0.870
JUN	1.465	0.101	1.030	0.919	0.968	0.903	0.761	0.503	2.042	<b>0.039</b>	1.349	0.578
JUNB	1.270	0.153	0.983	0.957	1.016	0.937	0.886	0.703	1.545	0.068	1.092	0.770
LIF	0.926	0.703	0.839	0.569	0.818	0.514	1.191	0.856	1.027	0.918	0.640	0.337
LTA	1.401	0.273	1.366	0.511	1.290	0.485	1.095	0.955	1.587	0.417	1.694	0.458
MAGI2	1.049	0.605	0.985	1.000	1.022	0.906	0.921	0.721	1.075	0.553	1.051	0.741
MAP2K1	0.993	0.912	1.027	0.802	1.015	0.877	1.005	0.976	0.971	0.736	1.051	0.735
MAP2K3	0.945	0.397	0.982	0.805	0.908	0.312	0.932	0.550	0.984	0.813	1.035	0.841
MAP2K4	0.995	0.923	1.043	0.659	1.021	0.836	1.032	0.762	0.970	0.718	1.055	0.718
MAP2K6	0.947	0.432	0.984	0.817	0.956	0.658	0.952	0.684	0.937	0.468	1.014	0.982
MAP2K7	1.012	0.866	0.995	0.952	1.019	0.843	1.009	0.964	1.004	0.965	0.982	0.900
MAP3K14	1.033	0.659	0.996	0.968	1.022	0.853	1.005	0.993	1.043	0.637	0.988	0.965
MAP3K5	0.960	0.503	1.035	0.793	0.950	0.566	0.967	0.761	0.971	0.675	1.105	0.515
MAP3K7	0.988	0.848	1.033	0.736	0.993	0.935	1.004	0.946	0.983	0.835	1.063	0.667

MAP3K8	0.989	0.989	1.093	0.415	0.961	0.832	1.227	0.198	1.020	0.761	0.979	0.997
MAPK1	0.976	0.734	1.025	0.785	0.977	0.817	0.994	0.976	0.974	0.781	1.058	0.671
MAPK10	0.720	0.081	0.669	0.135	0.953	0.847	1.023	1.000	0.579	<b>0.046</b>	0.444	<b>0.045</b>
MAPK11	1.014	0.912	1.006	1.000	1.076	0.539	1.209	0.165	0.956	0.662	0.834	0.290
MAPK12	1.028	0.701	0.924	0.508	1.082	0.482	0.952	0.749	0.975	0.869	0.895	0.532
MAPK13	1.041	0.548	0.995	0.996	0.990	0.977	0.998	0.984	1.092	0.355	0.992	0.983
MAPK14	0.972	0.621	1.047	0.689	1.010	0.941	0.985	0.931	0.936	0.400	1.110	0.544
MAPK3	1.026	0.762	1.031	0.769	0.978	0.787	0.959	0.718	1.074	0.474	1.103	0.465
MAPK8	0.982	0.818	0.982	0.862	1.001	0.970	1.001	0.991	0.963	0.692	0.963	0.793
MAPK9	0.991	0.917	0.967	0.727	0.970	0.779	0.940	0.609	1.012	0.898	0.994	0.977
MLKL	0.933	0.308	1.002	0.933	1.018	0.899	0.968	0.735	0.853	0.085	1.039	0.843
MMP14	1.025	0.868	1.124	0.427	1.045	0.747	1.078	0.601	1.005	0.952	1.167	0.556
MMP9	1.095	0.801	1.227	0.712	0.751	0.564	0.853	0.675	1.511	0.467	1.454	0.638
NFKB1	0.995	0.908	1.051	0.633	1.034	0.768	1.066	0.601	0.957	0.628	1.037	0.831
NFKBIA	1.214	0.153	1.043	0.797	1.185	0.399	1.221	0.553	1.239	0.248	0.883	0.787
NOD2	0.942	0.340	1.077	0.513	1.031	0.834	1.074	0.589	0.858	0.135	1.081	0.685
PGAM5	0.995	0.970	1.004	0.954	0.997	0.994	1.004	0.957	0.992	0.960	1.005	0.970
PIK3CA	1.012	0.886	1.002	0.985	1.007	0.953	1.006	0.973	1.016	0.879	0.999	0.999
PIK3CB	0.973	0.638	1.044	0.707	0.967	0.705	0.942	0.627	0.980	0.749	1.157	0.319
PIK3CD	0.970	0.671	0.952	0.607	0.966	0.731	0.922	0.503	0.973	0.772	0.982	0.901
PIK3CG	0.942	0.396	0.973	0.759	0.950	0.614	0.931	0.637	0.935	0.446	1.018	0.946
PIK3R1	0.994	0.993	0.974	0.820	0.976	0.798	1.055	0.686	1.011	0.810	0.898	0.473
PIK3R2	1.374	0.614	0.522	0.528	2.174	0.509	0.320	0.719	0.926	1.000	0.594	0.730
PIK3R2	0.987	0.953	0.889	0.616	0.966	0.918	0.992	1.000	1.008	1.000	0.784	0.482
PIK3R3	0.920	0.349	0.910	0.501	0.872	0.311	0.784	0.311	0.972	0.808	1.065	0.716
PIK3R5	1.023	0.730	1.041	0.664	1.040	0.682	1.088	0.505	1.006	0.927	0.997	0.986
PTGS2	0.815	0.583	0.557	0.202	0.472	0.166	0.598	0.306	1.262	0.564	0.528	0.399
RELA	0.978	0.773	0.960	0.662	0.979	0.837	0.998	0.961	0.977	0.822	0.924	0.559
RIPK1	0.943	0.413	0.979	0.816	0.949	0.596	0.940	0.619	0.937	0.490	1.019	0.906
RIPK3	0.949	0.443	0.962	0.614	0.968	0.734	0.944	0.607	0.929	0.411	0.981	0.809
RPS6KA4	1.025	0.767	1.074	0.477	1.021	0.861	0.978	0.857	1.029	0.793	1.175	0.243
RPS6KA5	1.013	0.862	1.032	0.717	0.989	0.934	1.022	0.817	1.038	0.724	1.044	0.750
SELE	1.342	0.424	1.152	0.874	1.181	0.762	0.789	0.998	1.508	0.539	2.039	0.623
SOCS3	0.964	0.913	0.835	0.664	0.983	0.988	0.866	0.857	0.951	0.905	0.802	0.629
TAB1	0.997	0.967	1.004	0.973	1.010	0.904	1.002	0.947	0.985	0.847	1.006	0.995
TAB2	1.050	0.485	1.013	0.859	1.004	0.975	0.998	0.993	1.094	0.324	1.028	0.803
TAB3	1.004	0.948	0.987	0.908	0.982	0.855	0.992	0.948	1.025	0.781	0.983	0.920

	<b>TNFAIP3</b>	1.180	0.367	1.122	0.673	1.153	0.598	1.580	0.282	1.203	0.471	0.785	0.709
	<b>TNFRSF1A</b>	0.933	0.331	1.041	0.790	0.977	0.792	0.917	0.548	0.891	0.289	1.169	0.418
	<b>TNFRSF1B</b>	0.988	0.800	1.055	0.604	0.999	0.949	0.964	0.761	0.976	0.747	1.157	0.291
	<b>TRADD</b>	1.057	0.432	1.054	0.584	1.063	0.529	1.075	0.548	1.051	0.590	1.035	0.804
	<b>TRAF1</b>	0.994	0.964	0.965	0.727	0.979	0.831	0.994	0.935	1.008	0.882	0.935	0.662
	<b>TRAF2</b>	0.997	0.959	0.997	0.979	1.006	0.990	1.051	0.712	0.988	0.921	0.943	0.684
	<b>TRAF3</b>	1.000	0.957	1.015	0.887	1.031	0.809	1.087	0.494	0.969	0.734	0.949	0.673
	<b>TRAF5</b>	1.031	0.650	0.985	0.909	1.005	0.953	1.064	0.599	1.058	0.535	0.912	0.516
	<b>VCAM1</b>	1.201	0.544	1.097	0.793	1.010	0.948	0.764	0.737	1.683	0.128	2.316	<b>0.049</b>
	<b>VEGFC</b>	0.872	0.542	0.705	0.260	1.144	0.880	0.709	0.574	0.697	0.161	0.703	0.360
<b>LT-BR</b>	<b>CCL13</b>	0.903	1.000	2.105	1.000	-	-	-	-	0.200	0.397	1.089	1.000
	<b>CCL19</b>	0.778	0.566	0.949	0.993	0.813	0.731	0.901	0.998	0.712	0.814	1.058	1.000
	<b>CCL21</b>	-	-	-	-	-	-	-	-	-	-	-	-
	<b>CHUK</b>	0.956	0.517	1.000	0.967	0.964	0.716	0.958	0.732	0.948	0.545	1.044	0.796
	<b>CXCL12</b>	1.443	0.582	1.516	0.720	2.988	0.389	0.454	0.998	0.977	0.991	1.744	0.655
	<b>ICAM1</b>	0.953	0.608	1.000	0.985	0.973	0.815	1.114	0.657	0.932	0.610	0.887	0.646
	<b>LTBR</b>	0.980	0.708	1.109	0.323	1.036	0.771	1.047	0.735	0.928	0.403	1.174	0.323
	<b>MAP3K14</b>	1.033	0.659	0.996	0.968	1.022	0.853	1.005	0.993	1.043	0.637	0.988	0.965
	<b>NFKB2</b>	1.020	0.834	1.026	0.818	1.083	0.464	1.090	0.507	0.960	0.649	0.964	0.771
	<b>RELA</b>	0.978	0.773	0.960	0.662	0.979	0.837	0.998	0.961	0.977	0.822	0.924	0.559
	<b>RELB</b>	1.013	0.901	1.049	0.622	1.079	0.492	1.150	0.257	0.952	0.606	0.957	0.738
	<b>TNFSF13B</b>	0.952	0.574	1.104	0.431	1.094	0.528	0.992	0.930	0.829	0.241	1.224	0.333

“-“ indicates that there were no reads among responders, non-responders or both.

**Supplementary Table 3:** Reactome Pathways up- or downregulated in non-responders compared to responders. All gene transcripts in the discovery cohort with a RPKM>0.5 (n=12,717) were analyzed for the presence of enriched Reactome pathways using a gene set enrichment analysis (GSEA) to identify pathways higher or lower expressed in non-responders compared to responders. All pathways were selected that showed a B&H adj. p-value < 0.20. Immune related pathways are marked with \*.

Pathway	Enrichment Score	Normalized Enrichment Score	NOM p-val	B&H adj. p-val	Total no. of genes in the pathway
Peptide chain elongation	0.716	2.333	<0.000001	<0.000001	83
Srp dependent cotranslational protein targeting to membrane	0.665	2.294	<0.000001	<0.000001	106
Nonsense mediated decay enhanced by the exon junction complex	0.654	2.270	<0.000001	<0.000001	104
3 utr mediated translational regulation	0.675	2.252	<0.000001	0.000200	102
*Influenza viral RNA transcription and replication	0.673	2.258	<0.000001	0.000250	99
*Interferon signaling	-0.606	-2.135	<0.000001	0.000498	129
*Interferon alpha beta signaling	-0.752	-2.225	<0.000001	0.000995	44
Translation	0.614	2.152	<0.000001	0.006207	141
*Interferon gamma signaling	-0.667	-2.012	<0.000001	0.012456	50
*Cytokine signaling in immune system	-0.522	-1.967	<0.000001	0.020246	225
Formation of fibrin clot clotting cascade	-0.804	-1.925	0.002037	0.029995	15
Rna pol I promoter opening	0.753	2.008	0.001901	0.047312	25
*Influenza life cycle	0.586	2.019	<0.000001	0.048181	132
Rna pol I transcription	0.673	1.963	0.003831	0.063036	43
Meiotic recombination	0.664	1.922	0.011719	0.078605	40
*Antigen presentation folding assembly and peptide loading of class II MHC	-0.751	-1.837	0.005952	0.085843	16
Response to elevated platelet cytosolic ca2	-0.583	-1.822	0.001961	0.087583	57
Activation of the mRNA upon binding of the cap binding complex and eifs and subsequent binding to 43s	0.622	1.885	0.001996	0.089904	54
Formation of the ternary complex and subsequently the 43s complex	0.635	1.855	0.001905	0.098261	46
Cell surface interactions at the vascular wall	-0.587	-1.793	0.002033	0.103925	58
RNA pol I RNA pol III and mitochondrial transcription	0.564	1.814	0.005964	0.113269	73
Packaging of telomere ends	0.709	1.798	0.007782	0.115306	23
Deposition of new cenpa containing nucleosomes at the centromere	0.641	1.774	0.015180	0.125367	32
Hemostasis	-0.452	-1.765	<0.000001	0.128020	303
Metabolism of mRNA	0.467	1.743	<0.000001	0.141775	199
Meiosis	0.545	1.702	0.010267	0.171493	63

**Supplementary Table 4:** Differentially expressed genes, per cohort and treatment. In the discovery cohort of 80 patients, gene transcripts were measured using RNA-sequencing and tested for the differential expression in (extreme) non-responders *versus* (extreme) responders. Genes with a minimal expression (RPKM > 0.5) were selected which showed a nominal p-value<0.05 in combination with a fold change (FC) < 0.67 or >1.5. The number of gene transcripts selected per subgroup were: all resp *versus* non- resp - 6; all excellent *versus* bad – 22; ADA resp *versus* non- resp – 31; ADA excellent *versus* bad – 102; ETN resp *versus* non- resp – 22; ETN excellent *versus* bad – 32. Shown are the 178 of unique genes selected and a “x” to indicate the subgroup(s) in which the gene was different between responders *versus* non-responders.

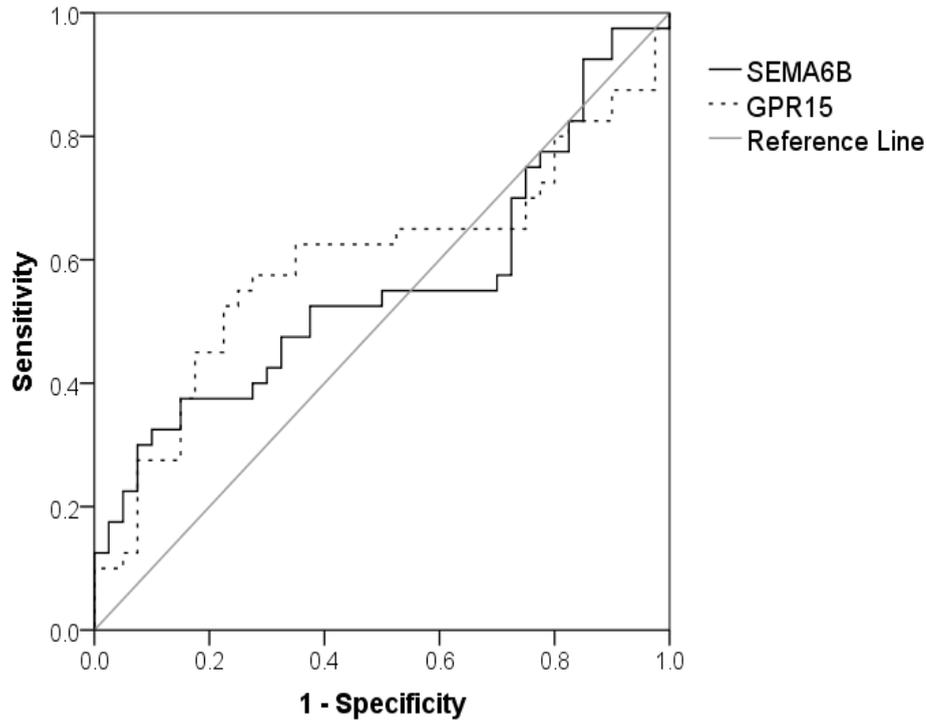
Gene transcript	All response	All extreme response	ADA response	ADA extreme response	ETN response	ETN extreme response	Number of subgroups (high to low)
SEMA6B	x	x	x	x			4
GPR15	x	x		x			3
KLK1		x	x	x			3
MARC2		x	x	x			3
RPL13P12		x	x	x			3
AC005154.8		x	x				2
AC006116.20			x	x			2
AC144530.1		x		x			2
ADORA3		x		x			2
BAIAP3			x	x			2
CLEC10A		x				x	2
COCH			x	x			2
DDX11L10					x	x	2
DDX11L2	x				x		2
ENHO		x				x	2
FADS2	x		x				2
GLRXP3		x				x	2
GPNMB		x				x	2
IGHV3-53				x			2
IGHV6-1				x			2
KCNIP2			x	x			2
MGAT3			x	x			2
NETO1			x	x			2
PARM1			x	x			2
RGS16		x		x			2
RP11-247113.3	x				x		2
RP11-848G14.5		x				x	2
SHISA4				x	x		2
SHISA8			x	x			2
SIGLEC6		x		x			2
WDR81		x				x	2
AC009963.3						x	1
AC234582.1			x				1
ATOX8						x	1
BACE2				x			1
BANK1				x			1
BCAT1						x	1
BLK				x			1
C1QC			x				1
C20orf195				x			1
CACNA1H						x	1

CCL3L3			x		1
CD19		x			1
CD1E				x	1
CD209				x	1
CD22		x			1
CD24		x			1
CD79A		x			1
CHAD		x			1
CHI3L2	x				1
COBLL1		x			1
COL19A1		x			1
COL4A3		x			1
CPNE5		x			1
CYP4F29P	x				1
DENND5B		x			1
DOK7		x			1
DPF3		x			1
DRAXIN	x				1
DTHD1	x				1
EBF1		x			1
EML6		x			1
ERAP2	x				1
FAM129C		x			1
FAM167A				x	1
FAM19A2		x			1
FAM209B				x	1
FAM21FP				x	1
FAM30A		x			1
FCRL1		x			1
FCRL2		x			1
FCRLA		x			1
FOLR3	x				1
FOS			x		1
FOSB			x		1
FPR3				x	1
GATM		x			1
GNG7		x			1
GPAT2				x	1
GPR84			x		1
GPS2P1		x			1
GSTM1		x			1
GYTL1B		x			1
HLA-A			x		1
HLA-B	x				1
HLA-C	x			x	1
IGHD		x			1
IGHG1		x			1
IGHM		x			1
IGHV1-69-2		x			1
IGHV3-13		x			1
IGHV3-15		x			1
IGHV3-43			x	x	1
IGHV3-74		x			1
IGHV4-34				x	1
IGHV4-39			x		1
IGHV6-1	x				1
IGKC		x			1

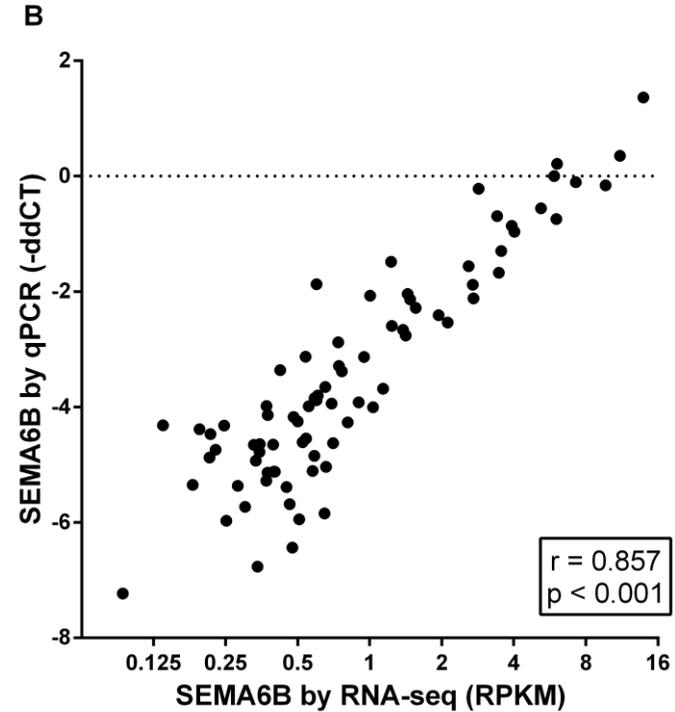
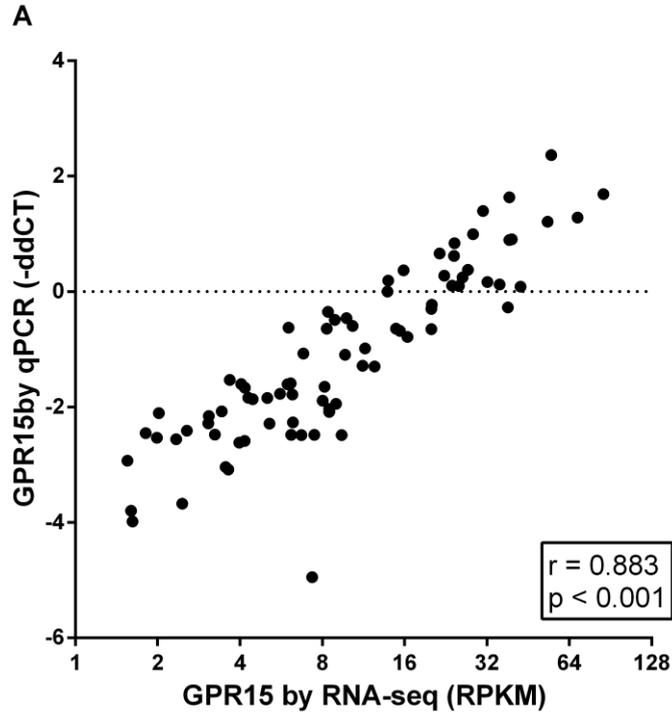
IGKV1-27		x			1
IGKV2-29		x			1
IGKV3-11		x			1
IGKV3-20		x			1
IGKV3D-20		x			1
IGLV9-49		x			1
IL9R		x			1
JUN			x		1
KCNH8				x	1
KLHL14		x			1
LAMA5		x			1
LAMC1		x			1
LARGE		x			1
LDHD			x		1
LILRB3				x	1
MACROD2		x			1
MME	x				1
MRC1				x	1
MS4A1		x			1
MTND4P12	x				1
NIPSNAP3B		x			1
NPM1P25		x			1
NRP1			x		1
OLFM1		x			1
OLIG1	x				1
ORM1		x			1
OSBPL10		x			1
P2RX5		x			1
PAWR		x			1
PAX5		x			1
PAX8			x		1
PHACTR1				x	1
PHOSPHO1	x				1
PID1				x	1
PKIG		x			1
PMEPA1				x	1
PNOC		x			1
POU2AF1		x			1
PXDC1		x			1
RAB6B		x			1
RALGPS2		x			1
RGL1		x			1
RP11-153M7.3			x		1
RP11-162A23.5	x				1
RP11-20024.1			x		1
RP11-22B23.1		x			1
RP11-343H5.4			x		1
RP11-475J5.4			x		1
RP11-497H16.7	x				1
RP11-767N6.2		x			1
RP11-793H13.10				x	1
RP11-927P21.7				x	1
RP13-128O4.3		x			1
RP4-765C7.2			x		1
RPH3A	x				1
RPH3AL		x			1
RPL13AP25		x			1

RPL9			x	1
RPS14P8			x	1
RPSAP53	x			1
SASH1	x			1
SGCE		x		1
SLAMF8				1
SLC2A5		x		1
SNX22		x		1
SPON1				1
SYNPO		x		1
TBC1D27		x		1
TCN1		x		1
THBS1	x			1
TNFRSF10C	x			1
TNFRSF13B		x		1
TNFRSF13C		x		1
TNIP3	x			1
TRBV6-4		x		1
TRBV6-6		x		1
VSIG4	x			1
ZC3HAV1L				1
ZNF608		x		1
ZNF860		x		1

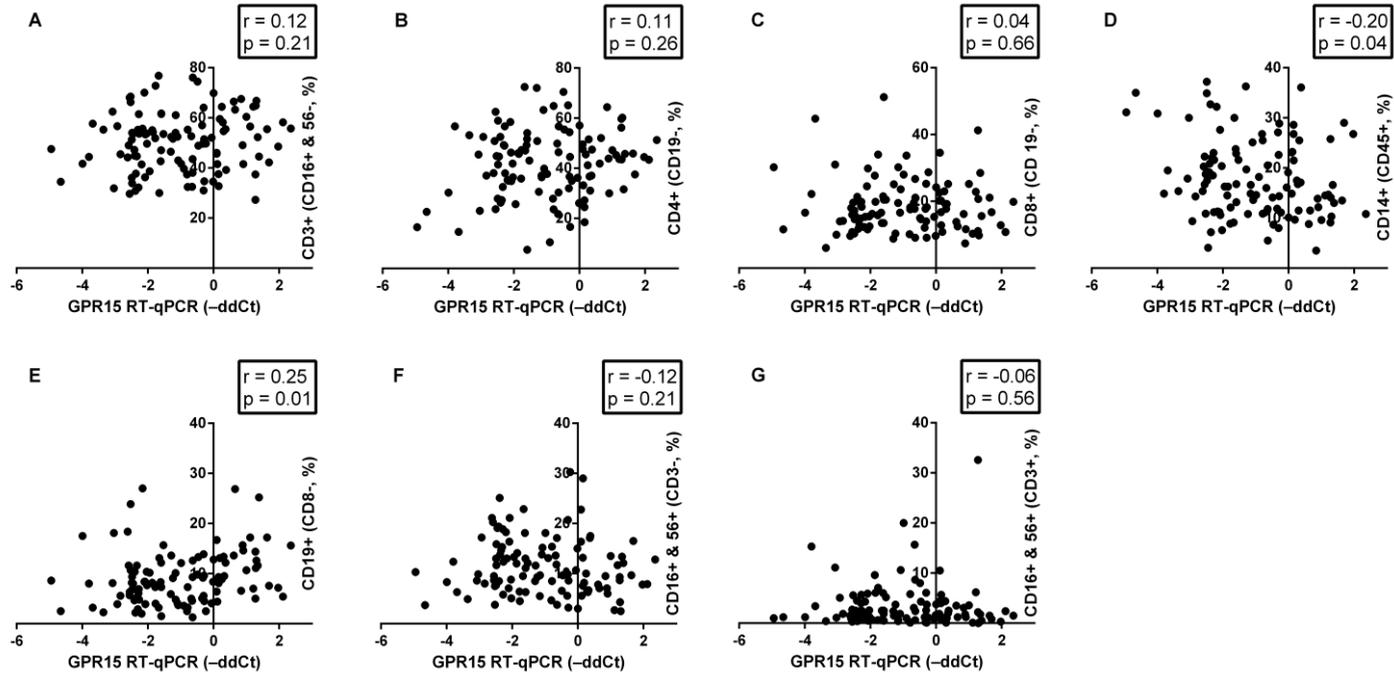
**Supplementary Figure 1:** Receiver operating characteristic curve for SEMA6B and GPR25 expression. Shown are the predictive abilities of GPR15 and SEMA6B (logtransformed value of the reads per kilobase per million mapped reads, RPKM) in distinguishing non-responders from responders, for different cut-offs. Both genes show a low sensitivity and high specificity, indicating they are especially suitably to identify non-responders with accuracy (high positive predictive value), yet less useful to pick up a response.



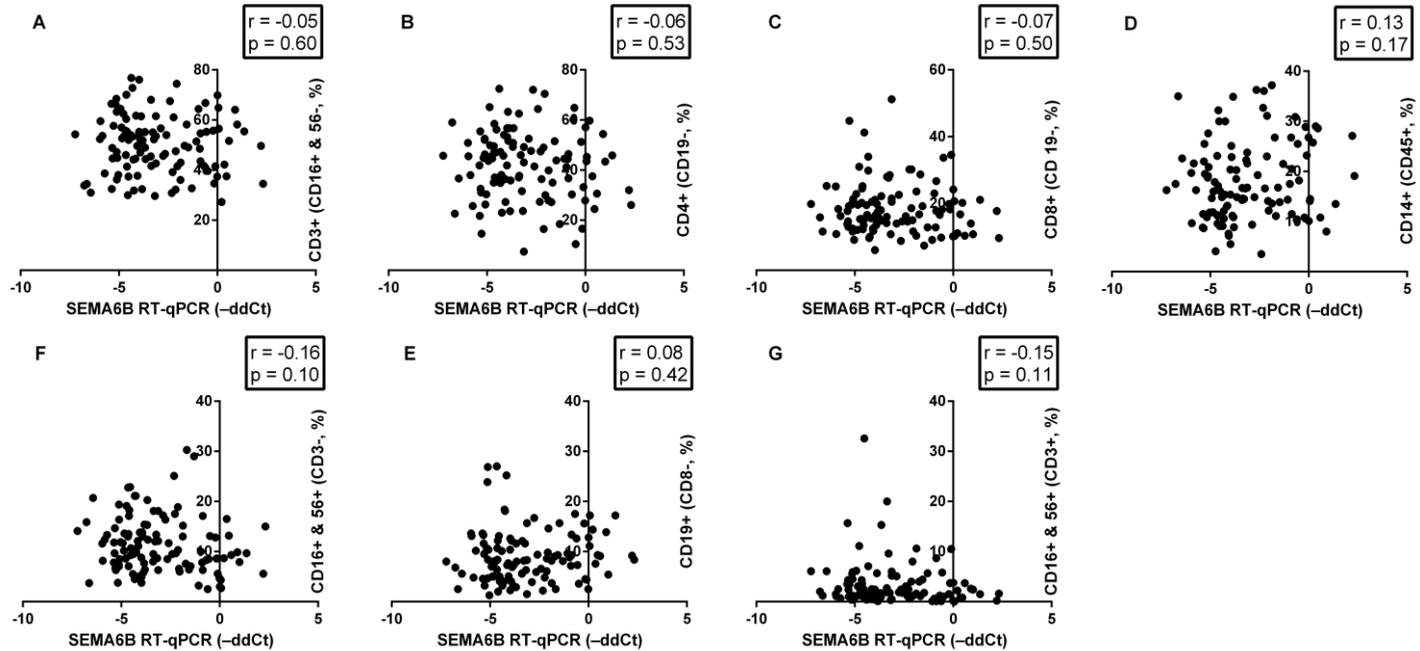
**Supplementary Figure 2:** Correlation between RNA-sequencing and single qPCR assay measurements of GPR15 and SEMA6B. The correlation between (A) GPR15 and (B) SEMA6B as measured by RNA-sequencing (reads per kilobase per million mapped reads, RPKM) and single qPCR assay (-ddCT value) was calculated by a Spearman's  $r$ , which followed a negative correlation due to the inverse relation of the CT value with expression. Both correlations were high, with small differences in favor of GPR15 which could be related to the (relatively) high expression.



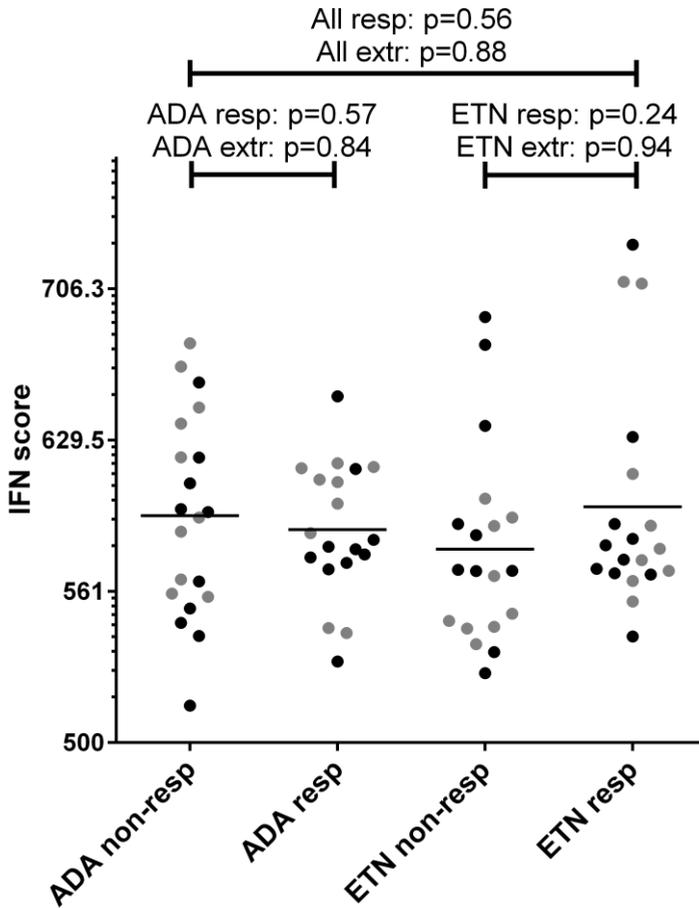
**Supplementary Figure 3:** Correlation between GPR15 expression (-ddCt) and PBMC cell-distribution. Immediately after PBMC isolation, the cells were stained for CD3, CD4, CD8, CD14, CD16+56, CD19, and CD45 and measured by flow cytometry (FACS), to identify the percentage of (A) T-cells, (B) Th-cells, (C) cytotoxic T-cells, (D) monocytes, (E) B-cells, (F) NK cells, and (G) NK T-cells. Each plot shows the correlation between GPR15 gene expression (-ddCt value) and the proportion of respective cell-subset. Correlations and p-values were calculated by a Spearman's  $r$ .



**Supplementary Figure 4:** Correlation between SEMA6B expression (-ddCt) and PBMC cell-distribution. Immediately after PBMC isolation, the cells were stained for CD3, CD4, CD8, CD14, CD16+56, CD19, and CD45 and measured by flow cytometry (FACS), to identify the percentage of (A) T-cells, (B) Th-cells, (C) cytotoxic T-cells, (D) monocytes, (E) B-cells, (F) NK cells, and (G) NK T-cells. Each plot shows the correlation between SEMA6B gene expression (-ddCt value) and the proportion of respective cell-subset. Correlations and p-values were calculated by a Spearman's  $r$ .



**Supplementary Figure 5: IFN-score in responders and non-responders.** Based on the study by Wright et al. (Rheumatology, 2015;54:188-93), an IFN-score was calculated by taking the sum of  $^2\log(\text{RPKM})$ -values of all selected IFN-related genes. Shown are the IFN-scores for the individual patients in the discovery cohort (n=80), the p-value as calculated by Mann-Whitney U and the geometric mean. The black dots represent the extreme (non-)responders and the grey dots the remaining (non-)responders. There was no significant difference in the IFN score between any of the subgroups.







## **Chapter 6**

### **CD4+ T cell mRNA evaluation by microarray and sequencing reveals no robust predictors of response to TNF-alpha inhibitors**

Bart V.J. Cuppen, Victor L. Jong, Arno N. Concepcion, Marita N. Nabibux, Floris P.J.G. Lafeber, Marinus J.C. Eijkemans, Ruth D.E. Fritsch-Stork, on behalf of all SRU investigators

## **ABSTRACT**

### **Objective**

Many studies have employed the use of messenger RNA (mRNA) to predict response to TNF $\alpha$ -inhibitors (TNFi) in rheumatoid arthritis (RA). Yet efforts to validate these targets have failed to show predictive abilities acceptable for clinical practice. We explored the transcriptome in different ways to identify and validate markers of response to TNFi therapy.

### **Methods**

The most extreme EULAR responders and non-responders to adalimumab and etanercept therapy after three months were selected from the observational BiOCURA cohort to compose three cohorts. On CD4+ cell enriched mRNA, collected before initiation of treatment, microarray was performed in the first cohort (n=39) and RNA-sequencing in the other two cohorts (n=16 and n=40). Differential expression was analyzed by MAANOVA in microarrays and by voom (limma-package) in RNA-seq. In each cohort, gene transcripts predicting response to either adalimumab or etanercept were selected based on a Benjamini & Hochberg FDR-corrected p-value <0.05, without or in combination with a minimal fold change (FC) of 2.0.

### **Results**

Although potential transcripts were identified in all three cohorts, there was no overlap in selection across the three cohorts. Two genes predicting response to etanercept (without FC as criterion) were overlapping between two of three cohorts and protein coding, of which IQSEC1 has been linked to the ARF6-pathway. In additional analyses, gene expression measured by microarray limitedly correlated with the same patients measured by RNA-seq, whereas the two RNA-seq cohorts highly correlated.

### **Conclusion**

This study was not able to identify predictive transcripts of TNFi response, consistent across multiple cohorts. Reasons for the irreproducibility lie in the technical variability of the microarray and the clinical heterogeneity in RA. Usage of RNA-seq and inclusion of factors to adjust for heterogeneity might increase the generalizability of results in future studies.

## INTRODUCTION

The introduction of tumour necrosis factor-alpha-inhibiting therapy (TNFi), has dramatically improved the treatment outcomes for rheumatoid arthritis (RA) patients. Still, not all patients experience an equally good response and 30-40% even hardly respond [1, 2]. As the clinical response may take several months to assess, non-responders are undertreated with the risk of irreversible joint damage. Ideally, non-response can be predicted before treatment initiation, so that other treatment options can be considered. In this respect, studying the messenger RNA (mRNA) at baseline might reveal gene transcripts that can predict response to TNFi therapy.

Microarrays changed the analysis of “one gene at a time” to genome-wide analyses [3], and has already been used for prediction of response to biological therapy in RA in at least twenty studies [4-27]. The results of these studies are heterogeneous and efforts to replicate models in external cohorts failed to reach predictive abilities acceptable for clinical practice [15, 24, 25]. Application of the newer techniques to study the transcriptome, such as by RNA-sequencing (RNA-seq), is limited so far. RNA-seq is considered to have a better reproducibility than microarray [28-30] thus possibly increasing the reproducibility of findings. To our knowledge, only two previous studies used RNA-seq for prediction of response to biological therapy in RA [31, 32]. Wright et al. [31] showed in neutrophil derived mRNA of twenty patients, that an interferon gene signature is able to predict response to TNFi therapy, whereas Folkerson et al. [32] selected targets from literature and narrowed this selection down to create a best performing model. However, both studies did not cross-validate their models in a new set of patients.

As replication is key in prognostic studies [33], we aimed to discover gene transcripts associated with TNFi response and subsequently assess their reproducibility across multiple cohorts, to assure the validity of possible prognostic markers for RA. Both microarray and RNA-sequencing were used for profiling of CD4+ T-cell enriched mRNA, as these cells are considered to be leading actors in the pathogenesis of RA [34].

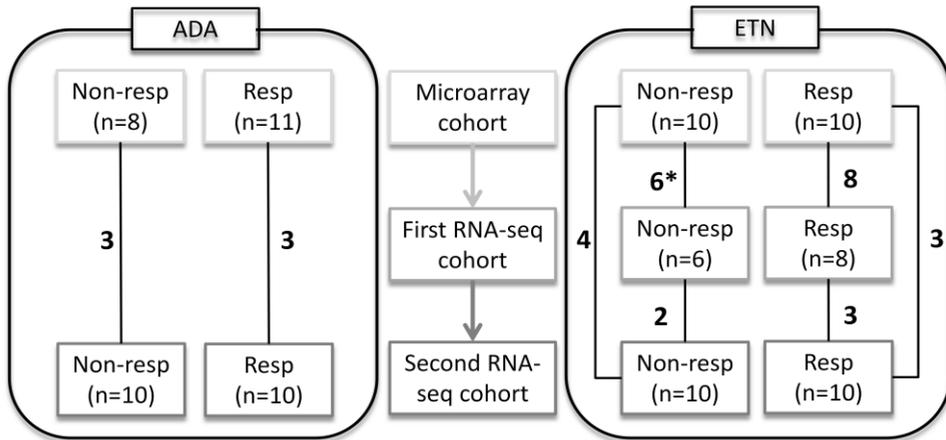
## **METHODS**

### **BiOCURA cohort**

Patients initiating adalimumab (ADA) and etanercept (ETN) therapy were selected from the “Biologicals and Outcome Compared and predicted Utrecht region in Rheumatoid Arthritis” (BiOCURA) study. BiOCURA is an observational cohort, in which RA patients eligible for biological treatment according to regular clinical practice were enrolled and followed after start of treatment, in one academic hospital and seven regional hospitals in the Netherlands. Re-inclusion after switching to a different biological treatment was possible, at which patients re-entered baseline again. The study was approved by the local ethics committee of the University Medical Center Utrecht and the institutional review boards of the participating centers (see Acknowledgements). Informed consent was obtained from each patient.

### **Patient selection**

In this study, the most extreme patients regarding clinical response to treatment were selected, as a clear distinctive clinical outcome might also give the clearest differences in gene expression at baseline, and would thus increase chances on finding relevant gene transcripts. The patients’ clinical response was first determined after three months of therapy using the EULAR response criteria, which is based on the absolute disease activity score (DAS28) and the change in DAS28 compared with baseline [35]. Then, patients were selected to form three cohorts, on which three separate gene expression experiments were performed: one microarray experiment, and a first and second RNA-seq experiment (Figure 1). First for the microarray cohort, eight non-responders and eleven responders for ADA and ten non-responders and ten responders for ETN, could be selected at the time. For the first RNA-seq cohort eight non-responders and eight responders for ETN were selected, fully overlapping with the microarray cohort in order to replicate these findings. For the subsequent RNA-seq cohort, again the ten best responders and ten non-responders for both ADA and ETN at that time were selected. Due to this selection method, there was a small overlap with the response groups of other cohorts.



**Figure 1:** Numbers of responders and non-responders (between brackets) with overlap (in bold) for the 2 cohorts of adalimumab (ADA) and three cohorts of etanercept (ETN) treated patients. Overlap in patients per experiment. Gene expression for these three cohorts was analyzed by microarray (performed in 2011, light grey) and RNA sequencing (first cohort in 2013, grey; second cohort in 2015, dark grey).

\* For the RNA-seq pilot originally 8 non-responders were selected, although 2 were excluded after analysis based on technical issues.

### Sample collection and processing

Blood was drawn shortly before treatment initiation in a Lithium-heparinized tube (BD Vacutainer®). Peripheral blood mononuclear cells (PBMCs) were isolated using Ficoll-Paque™ (GE healthcare Life Sciences, Little Chalfont, United Kingdom). The retrieved PBMCs were subsequently sorted using Magnetic Cell Isolation and Cell Separation (autoMACS Pro 360; Miltenyi Biotec, Bergisch Gladbach, Germany) by negative selection, using the CD4+ T Cell Isolation Kit, human (Miltenyi Biotec). FACS flow cytometry was performed to determine the purity of the CD4+ cells. RNA was extracted using the RNeasy Mini Kit (Qiagen) and tested for quantity and purity by Nanodrop (Thermo Scientific) and quality by RIN-value with Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA). All samples were aliquoted and stored per 1.5-2.0 µg at -80°C until use. All these and subsequent analyses were performed with well validated standard procedures. After applying these tests, the samples of the selected patients were considered to have a sufficient CD4+ T-cell purity (>80%, with one exception of 67%), RNA quality (RIN>8.0,

except for three samples with a value still over 7.1) and abundant quantity for all subsequent analyses.

### **Microarray**

Starting with 1.5 µg of RNA from each sample, cDNA synthesis, cRNA amplification, labeling, quantification, quality control and fragmentation were performed with an automated system (Caliper Life Sciences NV/SA, Belgium), all as previously described in detail [36]. In short, all patient RNA was labeled with Cy5, whereas reference cRNA (universal human reference RNA; Stratagene) was labeled with Cy3. Microarrays used were human whole genome gene expression microarrays V1 (Agilent, Belgium) representing 41,000 *Homo sapiens* 60-mer probes in a 4x44K layout. Microarray hybridization and washing was with a HS4800PRO system with QuadChambers (Tecan, Benelux) using 1000ng, 1-2% Cy5/Cy3 labeled cRNA per channel. Slides were scanned on an Agilent G2565BA scanner at 100% laser power, 30% PMT. After automated data extraction using Imogene 8.0 (BioDiscovery), mean spot-intensities were normalized using a non-parametric regression method (LOESS) [37]. Data were further analyzed by MAANOVA [38], in which p-values were determined by a permutation F2-test, and residuals were shuffled 10,000 times globally. All steps in the microarray experiment were performed by a specialized research facility within the University Medical Center Utrecht (Holstege laboratory).

### **RNA sequencing**

For analysis of the first and second RNA-seq cohort, the TruSeq® Stranded Total RNA kit (Illumina) was used for sample preparation. Starting with 1µg of CD4+ derived RNA, the low sample protocol was followed according to the manufacturer's instructions. The protocol includes the removal of rRNA with biotinylated target specific oligonucleotides, purification, fragmentation, priming with random hexameres, adding single "A" nucleotides to the 3' end, ligating multiple indexing adapters to the end of the ds cDNA and the enrichment of the DNA with PCR. The end product was tested for quantity and quality using the Qubit (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and Agilent 2100 Bioanalyzer. Sequencing was performed on an Illumina NextSeq500

sequencer (Illumina, San Diego, USA). For the first RNA-seq cohort (less deep), a 50bp single end run was performed, and the second RNA-seq cohort (more deep) three runs of 75bp single end. After FastQC quality control, aligning was performed against reference genome (GRCh37/Hg19), after which the read counts were normalized using DESeq [39]. On these normalized read a differential expression analysis was performed, for which the function “voom” [40] of the limma-package [41] was used in R version 3.2.4. In the first cohort RNA-seq, two non-responders were excluded for technical reasons (clear abnormalities in read counts). All experiments were directly supervised or performed by the Utrecht Sequencing Facility, with the only exception being the differential expression analysis.

### **Statistical analyses**

Two cut-offs were applied to select differentially expressed transcripts between (extreme) responders *versus* non-responders per subgroup: First, a Benjamini and Hochberg false discovery rate (B&H FDR) -corrected p-value <0.05 and second, a B&H FDR corrected p-value <0.05 in combination with a fold change (FC) <0.5 or >2.0. These selection criteria for potential transcripts are referred to as “p-value only” and “p-value and FC”, respectively. The overlap in selected genes per cohort, treatment, and applied selection criterion, was then investigated using Venny [42]. The functional relevance of the selected genes was further investigated by functional pathways enrichment analyses using the online available Toppgene software [43]. Test-retest analyses for patients overlapping among multiple cohorts were performed in SPSS version 21.0 (SPSS Inc., Chicago, USA), using the plot function and by calculation of the Spearman correlation ( $r$ ).

## **RESULTS**

Baseline characteristics for the subjects in the microarray cohort, and the first and second RNA seq cohort are shown in Table 1a and Table 1b, respectively. There were no statistically significant differences in clinical parameters between responders and non-responders, with the only exception of a higher C-reactive protein (CRP) in future responders compared with non-responders to ETN (median of 10.0 *versus* 3.7 resp.,

p=0.023) in the microarray cohort, and a more frequent use of concomitant glucocorticoids (GCs) in non-responders compared with responders to ADA (60% *versus* 0%, p=0.011) in the second RNA-seq cohort, with non-significant trends in most other cohorts.

**Table 1a:** Baseline characteristics of non-responders *versus* responders in the microarray cohort. The choice of the used test was based on the distribution of the parameter: independent sample t-test (normally distributed variables), Mann-Whitney U (non-normally distributed variables), Fisher exact (2x2) or chi-square (>2x2).

Clinical characteristic	microarray cohort - ADA			microarray cohort - ETN		
	Non-resp (n=8)	Resp (n=11)	p-value	Non-resp (n=10)	Resp (n=10)	p-value
Gender, females, n	7	6	0.117	8	8	1.000
Age, mean (sd)	54.8 (12.9)	57.9 (11.5)	0.591	58.8 (9.4)	54.8 (9.6)	0.360
Smoking currently, n	2	2	1.000	4	1	0.303
Alcohol >7 units/week, n	0	4	0.103	2	1	1.000
BMI, mean (sd)	27.9 (6.2)	24.6 (2.1)	0.195	27.0 (4.3)	31.2 (6.3)	0.096
RF, positive, n	4	9	0.319	8	7	1.000
ACPA, positive, n	5	9	0.603	7	8	1.000
CRP, mean (sd)*	5.0 (2.3-8.8)	9.0 (2.0-15.0)	0.310	3.7 (1.8-5.0)	10.0 (5.03-18.0)	<b>0.023</b>
Baseline DAS28, mean (sd)	3.9 (0.8)	4.4 (1.1)	0.294	3.9 (1.5)	4.5 (0.7)	0.243
No. of previous bDMARDs			1.000			1.000
0 (naïve), n	7	10		8	7	
1, n	1	1		2	1	
2, n	0	0		0	0	
Concomit HCQ, n	2	3	1.000	6	2	0.170
Concomit MTX, n	1	1	1.000	8	9	1.000
Concomit SZP, n	6	10	0.546	2	1	1.000
Concomit GCs, n	1	0	0.421	4	3	1.000

ACPA, anti-citrullinated protein antibody; bDMARDs, biological disease modifying antirheumatic drugs; BMI, body mass index; DAS28, disease activity score based on 28 joint count; GCs, glucocorticoids; HCQ, hydroxychloroquine; ESR, erythrocyte sedimentation rate; IQR, interquartile range; MTX, methotrexate; (Non-)Resp, (Non-)responder; RF, rheumatoid factor; SD, standard deviation; SJC, 28-swollen joint count; TJC, 28-tender joint count; VAS-GH, 100mm visual analogue scale on general health.

**Table 1b:** Baseline characteristics of non-responders *versus* responders in the first and second RNA-seq cohorts. The choice of the used test was based on the distribution of the parameter: independent sample t-test (normally distributed variables), Mann-Whitney U (non-normally distributed variables), Fisher exact (2x2) or chi-square (>2x2). For the RNA-seq pilot originally 8 non-responders were selected, although 2 were excluded after analysis based on technical issues (both with purity >80% and RIN>8.0).

Clinical characteristic	First RNA seq cohort – ETN			Second RNA seq cohort - ADA			Second RNA seq cohort – ETN		
	Non-resp (n=6)	Resp (n=8)	p-value	Non-resp (n=10)	Resp (n=10)	p-value	Non-resp (n=10)	Resp (n=10)	p-value
Gender, females, n	5	6	1.000	7	8	1.000	9	6	0.303
Age, mean (sd)	60.2 (10.0)	54.9 (10.6)	0.356	52.2 (12.3)	52.2 (13.2)	0.997	56.8 (10.3)	51.3 (11.0)	0.264
Smoking currently, n	1	0	0.429	6	1	0.057	3	3	1.000
Alcohol >7 units/week, n	1	1	1.000	1	1	1.000	1	3	0.582
BMI, mean (sd)	27.6 (4.8)	32.3 (6.6)	0.165	26.7 (5.1)	23.3 (3.7)	0.104	28.1 (5.3)	31.1 (7.2)	0.315
RF, positive, n	6	6	0.473	7	8	1.000	6	8	0.628
ACPA, positive, n	3	7	0.245	7	8	1.000	6	9	0.303
CRP median (IQR)*	4.0 (3.0-9.0)	11.5 (4.5-24.0)	0.181	10.0 (5.0-12.0)	5.0 (3.5-11.3)	0.400	5.5 (2.0-12.8)	8.0 (3.0-28.0)	0.400
Baseline DAS28, mean (sd)	3.9 (1.4)	4.4 (0.8)	0.403	4.5 (1.3)	4.5 (1.0)	0.895	4.9 (1.3)	4.7 (0.8)	0.649
No. of previous bDMARDs			1.000			0.264			0.587
0 (naïve),n	5	7		7	9		7	8	
1, n	1	1		3	1		2	2	
2, n	0	0		0	0		1	0	
Concomit HCQ, n	3	2	0.580	4	1	0.303	4	4	1.000
Concomit MTX, n	5	7	1.000	6	8	0.628	6	9	0.303
Concomit SZP, n	1	1	1.000	1	2	1.000	1	1	1.000
Concomit GCs, n	0	2	0.473	6	0	<b>0.011</b>	5	2	0.350

ACPA, anti-citrullinated protein antibody; bDMARDs, biological disease modifying antirheumatic drugs; BMI, body mass index; DAS28, disease activity score based on 28 joint count; GCs, glucocorticoids; HCQ, hydroxychloroquine; ESR, erythrocyte sedimentation rate; IQR, interquartile range; MTX, methotrexate; (Non-)Resp, (Non-)responder; RF, rheumatoid factor; SD, standard deviation; SJC, 28-swollen joint count; TJC, 28-tender joint count; VAS-GH, 100mm visual analogue scale on general health.

### Large range in number of found transcripts per cohort, treatment and criterion

Per cohort, treatment and selection criterion, a variable number of gene transcripts was found differentially expressed between non-responders and responders (Table 3). In the microarray cohort, only six gene transcripts were significantly different between responders and non-responders to ADA, whereas 206 probes differed between responders and non-responders in ETN treated patients. However, of these 206 targets only one transcript remained after applying the additional cut-off for the FC (<0.5 or >2.0). This gene is known for synaptic functions in neurons (i.e. leucine rich repeat neuronal 3, LRRN3). In the first RNA-seq cohort, performed on ETN treated patients fully overlapping with the microarray cohort, 18 (p-value only) and 15 (p-value and FC) transcripts were found to predict response. Finally, in the second RNA-seq cohort three gene transcripts were found for ADA (p-value only and p-value and FC) and 19,470 transcripts for ETN (p-value only), the latter of which corresponding with almost a third of all identified transcripts (63,677 in total). The application of the FC brought the number for ETN down to 458.

**Table 3:** Number of found genes, per cohort and treatment. The selection was either based on a Benjamini & Hochberg false discovery rate (B&H FDR) corrected p-value<0.05 only (“p-value only”), or B&H FDR corrected p-value in combination with a fold change (FC) < 0.5 or >2.0 (“p-value and FC”).

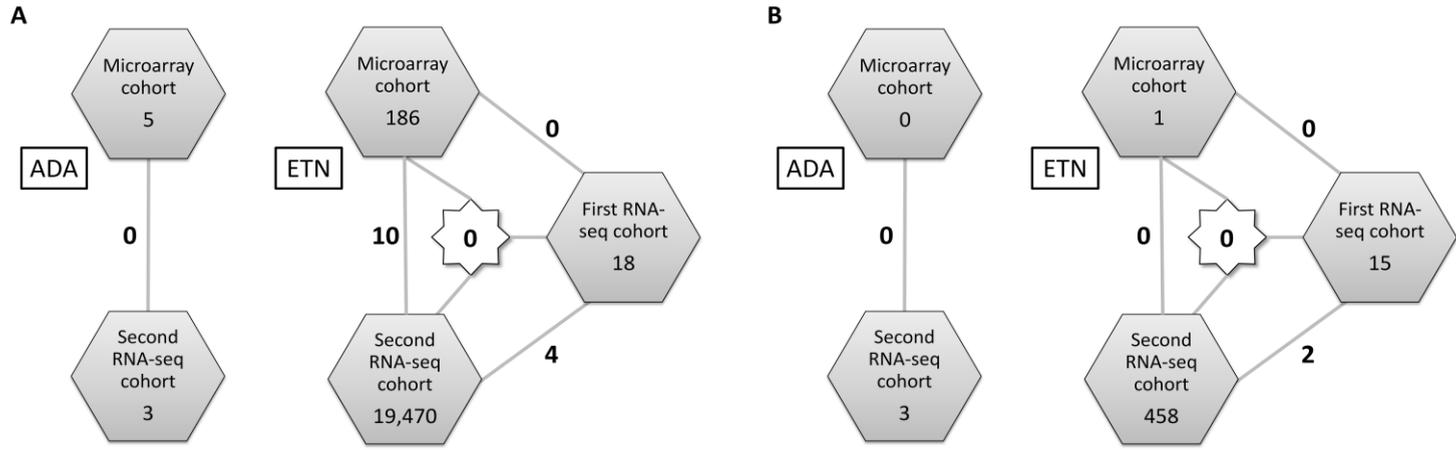
Treatment	Study cohort + technique	p-value only	p-value and FC
ADA	Microarray cohort	5 (6*)	0
ADA	Second RNA-seq cohort	3	3
ETN	Microarray cohort	206 (186*)	1
ETN	First RNA-seq cohort	18	15
ETN	Second RNA-seq cohort	19470	458

\*Number of probes originally selected based on criterion, as duplications occurred in case  $\geq 2$  probes on the Microarray chip are mapped to the same gene.

### Limited overlap in found genes among cohorts

The overlap between the three cohorts in selected gene transcripts was investigated: no overlapping genes were found in both ADA cohorts or in all three ETN cohorts, regardless

of selection criteria (Figure 2). Additionally, only fourteen (p-val only) and two (p-val and FC) genes were found in two of the three ETN cohorts. Moreover, not only was there limited overlap in genes between the different techniques (microarray cohort and the other cohorts), but also between the cohorts measured using the same technique (first and second RNA-seq cohort), all on top of the partial overlap in patient selection. Although the selected genes were not considered to be completely robust, they might still constitute interesting predictors, especially those overlapping among two of three cohorts, and were therefore further investigated.



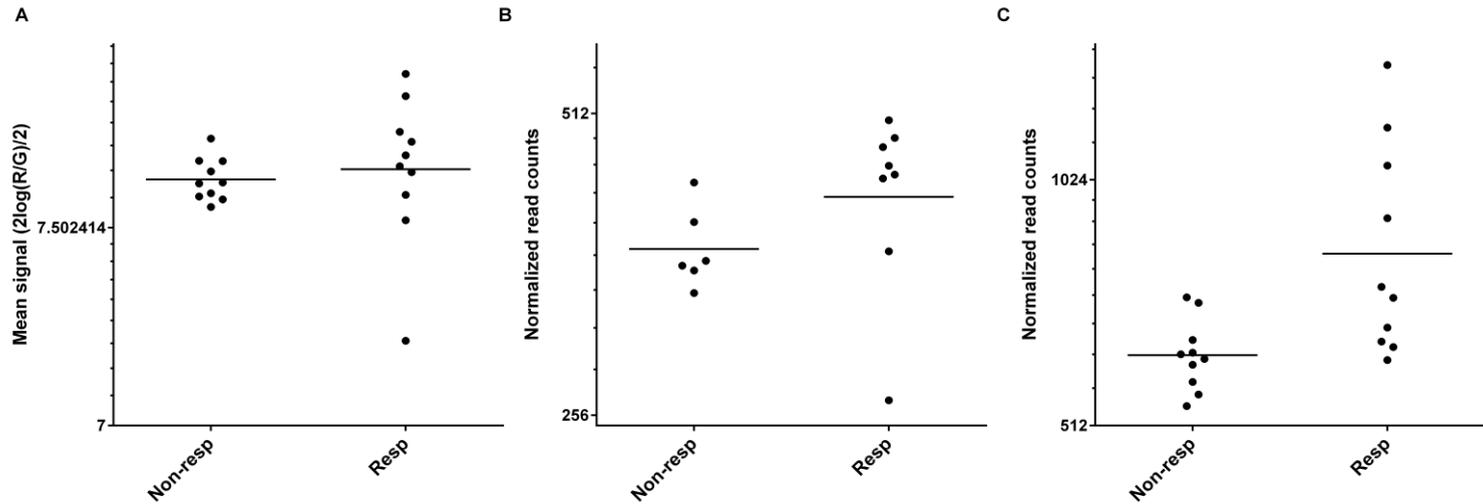
**Figure 2:** Overlap in genes among cohorts. The selected genes were compared across the cohorts as measured by either microarray or RNA-seq. For **(A)** Selected genes based on p-value < 0.05 and **(B)** selected genes based on p-value in combination with a FC < 0.5 or > 2.0, the total number of selected gene transcripts is shown in the hexagons whereas the overlap between cohorts is shown in the numbers next to the connecting lines.

### **IQSEC1 (part of in ARF6 pathway) found, yet complete ARF6 pathway not predictive**

A pathway enrichment analysis was performed on all selected transcripts in all cohorts and among those overlapping between two of three cohorts (brief results in Table 4). Interesting biological functions and pathways were represented within some selections, such as in the 458 genes from the second RNA-seq study in ETN (p-value and FC) that contained three of the seven genes from the “Antigen processing/presentation exogenous lipid antigen via MHC class Ib” pathway (GO:0048007, B&H adj. p-val = 7.60E-03). Among all fourteen genes that overlapped between two of the three ETN cohorts, only two genes could be annotated to known biologically functional regions of the genome. One of these genes is IQ Motif and Sec7 Domain 1 (IQSEC1), which is a component of the ARF6 pathway (1 of 34 genes in total, p=6.23E-03). When retrieving the individual expression of this gene in all datasets, significantly higher read counts of IQSEC1 were seen in future responders to ETN therapy in both RNA-seq cohorts, but not in the microarray cohort (Figure 3). Since the ARF6 pathway was considered potentially interesting, the FCs and p-values for other genes in the ARF6 pathway were retrieved from all experiments (Table 5). We saw no other up- or down-regulated gene transcripts, however EGF, ITGA2B, ITGB3 were down-regulated to some extent in ETN responders, with a FC<0.5 in the first and second RNA-seq cohorts (not significant).

**Table 4:** Results from pathways analyses. The found genes from each experiment were put in the online pathway enrichment analyses software Toppgene. Shown are the highest inflammatory related GO: biological processes or pathways from the analysis, and in case no inflammatory pathways emerged the highest placed pathway was shown. Genes found in two of the respective three ETN cohorts (14 based on  $p < 0.05$  and 2 on  $p < 0.05$  and  $FC > 2$ ), were considered to be biologically replicated to some extent. Only two transcripts in ETN patients could be annotated to any protein coding gene: BTF3 (basic transcription factor 3, overlapping between microarray and second RNA-seq cohort) and IQSEC1 (IQ Motif And Sec7 Domain 1, overlapping between the first and second RNA-seq cohort). IQSEC1 was related to the ARF6 pathway ( $p = 6.23E-03$ ).

Treatment	Study cohort	Selection criterion	Selected genes	Annotated genes	Pathway	q val FDR B&H	No. of genes /Total in pathway
ADA	Microarray cohort	$P < 0.05$	5	3	TGF-beta receptor signaling	2.46E-02	1/51
		$P < 0.05 + FC > 2.0$	0	0	-	-	-
ADA	Second RNA-seq cohort	$P < 0.05$	3	2	Neuropeptide signaling pathway	1.15E-02	1/107
		$P < 0.05 + FC > 2.0$	3	2	Neuropeptide signaling pathway	1.15E-02	1/107
ADA	Overlapping	$P < 0.05$	0	0	-	-	-
		$P < 0.05 + FC > 2.0$	0	0	-	-	-
ETN	Microarray cohort	$P < 0.05$	186	92	Viral mRNA translation	2.06E-56	38/140
		$P < 0.05 + FC > 2.0$	1	1	-	-	-
ETN	First RNA-seq cohort	$P < 0.05$	18	8	Multidrug resistance protein mediated transport	1.47E-02	1/1
		$P < 0.05 + FC > 2.0$	15	6	Pos. regulation of antigen processing and presentation via MHC class I (GO:0002591)	2.63E-02	1/6
ETN	Second RNA-seq cohort	$P < 0.05$	19470	4920	Antigen processing/presentation	2.45E-08	30/82
		$P < 0.05 + FC > 2.0$	458	224	Antigen processing/presentation, exogenous lipid antigen via MHC class Ib (GO:0048007)	7.60E-03	3/7
ETN	Overlapping	$P < 0.05$	14	2	ARF6 signaling events	6.23E-03	1/34
		$P < 0.05 + FC > 2.0$	2	0	-	-	-



**Figure 3:** IQSEC1 expression at baseline in ETN initiating patients. Shown are the normalized read counts for IQSEC1 for responders *versus* non-responders to ETN in the (A) Microarray cohort (i.e. mean of 5 probes targeting IQSEC1, all  $p=1.000$ ) (B) first RNA-seq cohort ( $p=0.042$ ) and (C) second RNA-seq cohort ( $p=0.033$ ). The horizontal bar indicates the geometric mean.

**Table 5:** ARF6 related genes. Because the IQSEC1 gene was overlapping (ETN pilot and full study, p-val<0.05), a gene list of the ARF6 signaling pathway, containing 34 genes, was derived from BioSystems [<http://www.ncbi.nlm.nih.gov/biosystems/138034>]. From each experiment these 34 genes were retrieved, in order to look for patterns within this pathway. The relative expression of responders compared to non-responders is shown as fold change (FC). P-values <0.05 and FC<0.5 or >2.0 were made bold.

Gene	Ensemble ID	ADA				ETN					
		Microarray cohort		Second RNA-seq cohort		Microarray cohort		First RNA-seq cohort		Second RNA-seq cohort	
		FC	P-val	FC	P-val	FC	P-val	FC	P-val	FC	P-val
ACAP1	ENSG00000072818	0.914	1.000	1.033	0.721	0.977	1.000	1.169	0.082	1.073	0.612
ACAP2	ENSG00000114331	1.000	1.000	1.062	0.547	1.106	0.932	0.957	0.447	0.945	0.352
ADAP1	ENSG00000105963	1.041	1.000	0.883	0.844	0.961	1.000	1.225	0.326	1.069	0.886
ADRB2	ENSG00000169252	1.319	1.000	1.161	0.816	1.093	1.000	0.921	0.865	0.912	0.893
AGTR1	ENSG00000144891	0.972	1.000	1.068	0.792	1.030	1.000	0.984	0.791	0.981	0.932
ARAP2	ENSG00000047365	1.081	1.000	1.023	0.943	0.995	1.000	1.077	0.645	0.961	0.763
ARF6	ENSG00000165527	1.117	1.000	0.929	0.648	1.029	1.000	1.016	0.828	0.983	0.909
ARRB2	ENSG00000141480	1.025	1.000	1.007	0.977	0.975	1.000	1.122	0.166	0.993	0.959
CYTH2	ENSG00000105443	0.914	1.000	1.207	0.547	0.934	1.000	1.199	0.082	1.096	0.207
CYTH3	ENSG00000008256	0.898	1.000	1.162	0.657	1.002	1.000	1.915	0.082	1.080	0.801
EFNA1	ENSG00000169242	1.042	1.000	1.375	0.633	1.141	0.887	1.090	0.862	1.986	0.071
EGF	ENSG00000138798	1.004	1.000	1.641	0.702	0.985	1.000	<b>0.312</b>	0.160	<b>0.316</b>	0.395
EGFR	ENSG00000146648	1.009	1.000	1.432	0.547	1.051	1.000	0.941	0.791	0.770	0.467
EPHA2	ENSG00000142627	0.999	1.000	0.932	0.977	1.045	1.000	1.691	0.324	0.530	0.261
FBXO8	ENSG00000164117	1.026	1.000	1.191	0.547	0.974	1.000	0.940	0.738	0.991	0.944
GIT1	ENSG00000108262	0.974	1.000	0.940	0.616	0.959	1.000	1.024	0.795	1.147	0.052
GNA11	ENSG00000088256	0.962	1.000	1.409	0.547	1.051	1.000	0.807	0.405	1.723	<b>0.033</b>
GNA14	ENSG00000156049	0.976	1.000	0.861	0.850	1.064	1.000	1.305	0.226	1.215	0.714
GNA15	ENSG00000060558	0.974	1.000	1.049	0.952	1.056	1.000	1.194	0.608	0.713	0.329
GNAQ	ENSG00000156052	1.033	1.000	1.049	0.702	1.034	1.000	0.988	0.884	0.867	0.091
GULP1	ENSG00000144366	1.013	1.000	0.895	0.869	1.128	0.641	1.027	0.878	1.039	0.942

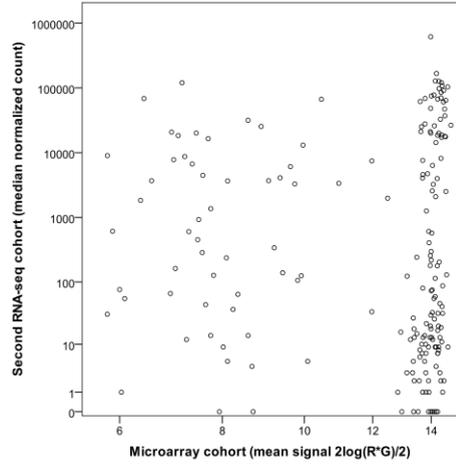
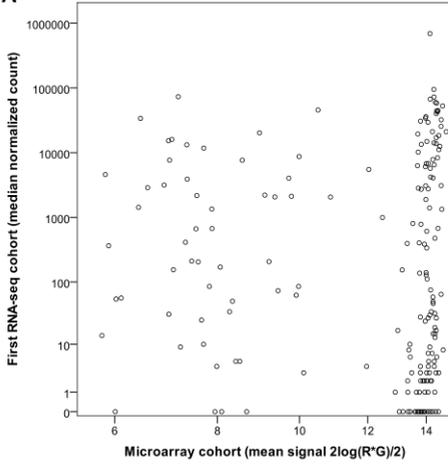
HGF	ENSG00000019991	1.085	1.000	1.365	0.787	1.100	0.998	1.303	0.741	0.644	0.478
IPCEF1	ENSG00000074706	0.892	1.000	0.950	0.773	0.917	1.000	0.892	0.082	1.085	0.629
IQSEC1	ENSG00000144711	1.059	1.000	0.909	0.547	1.072	1.000	1.276	<b>0.042</b>	1.249	<b>0.033</b>
ITGA2B	ENSG00000005961	1.025	1.000	1.017	0.987	0.886	1.000	<b>0.386</b>	0.255	<b>0.379</b>	0.348
ITGB3	ENSG00000259207	1.091	1.000	0.729	0.810	0.868	1.000	<b>0.386</b>	0.438	<b>0.185</b>	0.281
KIF13B	ENSG00000197892	0.947	1.000	1.012	0.940	1.111	0.887	1.123	0.159	1.148	0.116
LHCGR	ENSG00000138039	0.994	1.000	0.973	0.946	0.998	1.000	0.957	0.791	1.097	0.543
MET	ENSG00000105976	0.930	1.000	0.695	0.615	1.141	0.314	<b>2.068</b>	0.082	0.761	0.682
NCK1	ENSG00000158092	1.048	1.000	1.044	0.606	1.062	1.000	0.934	0.224	1.061	0.604
PXN	ENSG00000089159	1.080	1.000	0.896	0.547	0.951	1.000	0.990	0.862	1.059	0.504
SRC	ENSG00000197122	0.971	1.000	1.254	0.683	1.050	1.000	0.854	0.778	0.660	0.527
TSHR	ENSG00000165409	1.019	1.000	0.551	0.547	0.932	1.000	0.545	0.281	0.719	0.552
USP6	ENSG00000129204	0.872	1.000	<b>2.617</b>	0.547	0.985	1.000	1.036	0.936	0.534	0.392

---

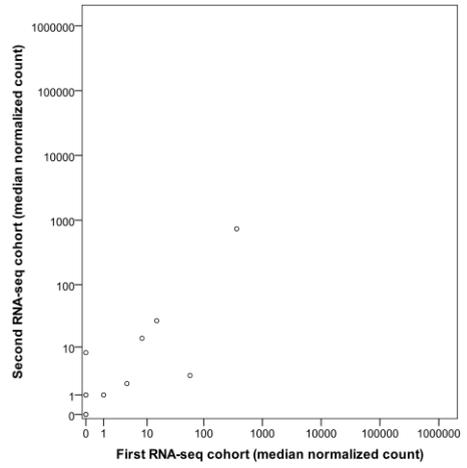
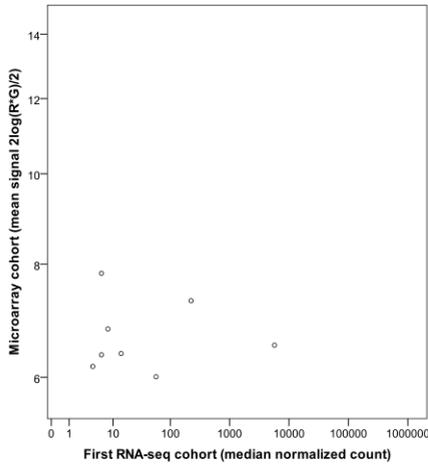
### **Gene expression as measured by microarray and RNA-seq weakly correlates**

Due to the very low overlap in selected genes, we additionally investigated the technical variability of all selected genes across the three cohorts. This was done by plotting and calculating the correlation in the expression of these genes between patients overlapping among cohorts (Figure 4). The plots showed a weak correlation between the expression in microarray and RNA-seq cohorts, regardless whether genes were found in the microarray cohort self ( $r=0.103$  and  $r=0.088$  with resp. first and second RNA-seq cohort), or found in the first or second RNA-seq cohort ( $r=0.156$  and  $r=0.271$  resp). The correlation between the first and second RNA-seq cohort was strong, with  $r=0.76$  and  $r=0.93$  for genes selected in the first and second RNA-seq cohort respectively. Additionally, the correlation in gene expression of all genes was investigated (up to 60,000 targets) (Supplementary Figure 1). These plots showed again a relatively weak correlation between the microarray and two RNA-seq cohorts ( $r=0.40$ ) and a strong correlation between the two RNA-seq studies ( $r=0.94$ ).

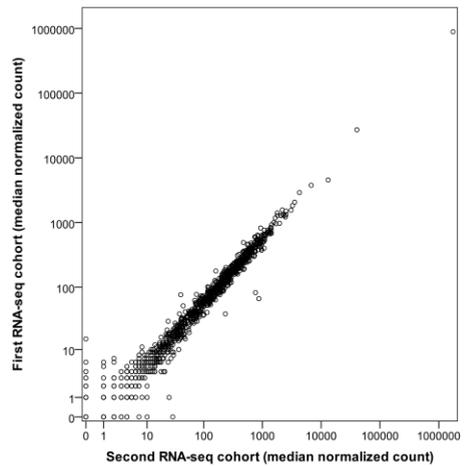
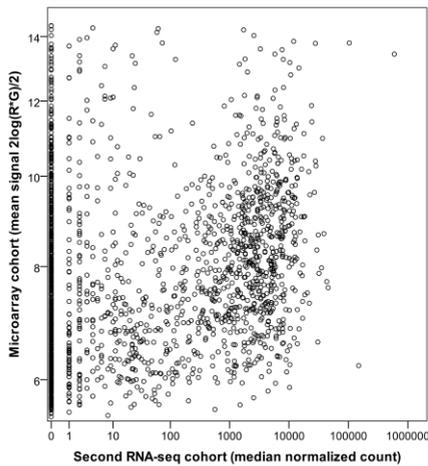
**A**



**B**



**C**



**Figure 4:** Correlation in expression levels of the found predictive genes, in similar patients across cohorts. One of the reason why the cross-validation between the three cohorts failed could be related to different expression levels as measured by the used technique. The expression of the differentially expressed genes (based on p-value) of patients overlapping between the two cohorts (Figure 1), are expected to correlate as it concerns the same patients and the same (relevant) genes, only another measurement. Expression levels were plotted for genes selected in the (A) microarray cohort (x-axis) and compared with their values in the 1<sup>st</sup> RNA-seq cohort ( $r=0.103$ ,  $p=0.156$ ) and 2<sup>nd</sup> RNA-seq cohort ( $r=0.088$ ,  $p=0.228$ ) (y-axis), (B) first RNA-seq cohort (x-axis) and compared with their values in the microarray ( $r=0.156$ ,  $p=0.713$ ) and 2<sup>nd</sup> RNA-seq cohort ( $r=0.763$ ,  $p<0.001$ ) (y-axis), and (C) second RNA-seq cohort (x-axis) and compared with their values in the microarray ( $r=0.271$ ,  $p<0.001$ ) and 1<sup>st</sup> RNA-seq cohort ( $r=0.931$ ,  $p<0.001$ ) (y-axis). These plots ideally show some sort of (non-parametric) agreement between the median expression values.

Of note, not all transcripts identified in the RNA-seq cohorts were probed in the microarray, which results in some missing values (left plot B and C). Also, for comparisons of the two RNA-seq experiments (right plot of B & C), the median normalized read counts are frequently overlapping (e.g.  $x=1, y=1$ ), giving the appearance of less values in the plot than are present and used in calculation of the correlation.

## DISCUSSION

Our study did not reveal gene transcripts in baseline CD4+ T-cell mRNA that consistently predict a clinical response to ADA or ETN. Although in all three cohorts the analysis revealed genes significantly differentially expressed between responders and non-responders and these genes were partly implicated in interesting pathways, no genes were overlapping between all cohorts, and only a limited number in two of three ETN cohorts, despite a partially overlapping patient selection. The differences in affected pathways between responders and non-responders and possible reasons for the absent overlap across cohorts are discussed below.

Several general immunologic pathways were found significantly overrepresented in the genes selected in the two RNA-seq cohorts, among many other pathways. The differential expression of genes involved these immunologic pathways, could therefore fit in the recently suggested hypothesis in which TNFi's (and non-TNFi's) target one final common pathway, namely pro-inflammatory cytokine production [44, 45]. Two aspects are of note: First, the results and p-values of any pathway enrichment analysis should be interpreted as "hypothesis generating", in this particular case also as they calculate

overrepresentation of genes compared to the background of the entire genome, whereas immunologic genes are generally higher expressed in CD4+ cells and more likely to be selected. Second, although these general immunologic pathways were selected, the found genes were sometimes all upregulated at baseline in responders to ETN (i.e. most of the 30 genes belonging to “Antigen processing/presentation”), yet sometimes all downregulated (the three genes belonging to “Antigen processing/presentation, exogenous lipid antigen via MHC class Ib”). Inherent to the hypothesis generating characteristic, other studies found pathways with widely varying biological processes involved, but usually also including general immunologic pathways [7, 10-12, 15, 18, 20, 22-24, 27, 31]. Whether or not general immunologic pathways are predictive of response remains therefore questionable.

Among the overlapping gene transcripts predicting response to ETN in both RNA-seq studies, IQSEC1 (IQ Motif And Sec7 Domain 1) was identified. The protein product of IQSEC1, GEP100, is known to activate ARF6 with high affinity [46], making the ARF6 pathway potentially interesting in the response to TNFi. The role of ARF6 has previously been implicated in effects of IL1b in sepsis [47] and a similar engagement in inflammatory cascade of TNF $\alpha$  was more recently described [48]. In our study, no up-/down-regulation of other genes present in the ARF6 signaling pathway was seen. This suggests that the influential mechanisms of IQSEC1 in TNFi response might be on proteomic level only, or that these are uncoupled from the rest of the ARF signaling pathway. The latter is not unlikely, given that GEP100 is known to be inducing apoptotic cell death in monocytes and even essential in TNF $\alpha$  induced apoptosis, possibly via independent mechanisms from ARF-activation [49]. As for the other gene overlapping two of the three ETN cohorts (i.e. microarray and second RNA-seq), is basal transcription factor 3 (BTF3), which is a transcription factor for RNA polymerase II. BTF3 is known to involve in apoptosis regulation and is considered to play a role in a range of different cancers and [50-52]. Although not consistently selected across cohorts, these two genes with apoptosis related functions might be interesting with relation to the pharmacological response to ETN.

The overlap between the three cohorts in predictive transcripts was limited. In fact, less overlap was seen between the ETN microarray cohort (186 genes) and the second RNA-seq cohort (19,470 genes, roughly one-third of all measured transcripts) than expected based on chance: instead of an overlap of 1 out of every 3 significant transcripts of the microarray (i.e.  $1/3 * 186 = 62$ ), only 10 genes were overlapping. Besides, even when prediction of treatment response is not possible and we are left to find general prognostic factors for worse outcome in RA (e.g. expression related to a baseline high disease activity or autoantibody status [53, 54]), we would still expect to find more overlap in CD4+ enriched mRNA that distinguishes responders from non-responders. Therefore, it is likely that other factors have played a role, such as the technical reproducibility (i.e. microarray *versus* RNA-seq) and the clinical heterogeneity in RA.

First some technical aspects could have had an influence. Although the mRNA quality before freezing was usually high, there was a time-interval up to 68 months for the samples before they were thawed (BiOCURA inclusion 2009 to-2015, microarray in 2011, and first and second RNA-seq in 2013 and 2015 respectively). Some decay and decrease in RIN-value in this time period is expected [55] and might have affected the results, especially in the latest performed experiment, i.e. the second RNA-seq cohort. However, the very high correlation between the first and second RNA-seq cohorts suggest that this influence is low. The other technical aspects is the difference in sequencing depth between both RNA-seq experiments. This is not expected to have a role in the irreproducibility, as normalization was performed using DESeq for both datasets and this would only explain why low abundant transcripts found in the second RNA-seq cohort (more deep) could not be replicated in the first RN-seq cohort (less deep), but not vice versa.

Second, the technical variability was larger than expected and may have introduced noise in our results, which subsequently could have contributed to the very limited overlap. Specifically, the correlation between the microarray and RNA-seq studies in absolute gene expression of all measured transcripts was unsatisfactory with an  $r$  of 0.4. These results clearly indicate that the gene expression as measured in the microarray is not directly comparable with gene expression by RNA-seq. In contrast, the two RNA-seq cohorts showed a high correlation. These results confirm the high reproducibility of RNA-seq in

literature [28-30], which makes RNA-seq probably the better choice for future biomarker discovery.

Lastly, we did not account for heterogeneity in RA that might interact in the relationship between gene expression and response. Although confounding is traditionally not considered an issue in prognostic research, it exists and is known to influence generalizability [56, 57]. In this particular case, correction for CRP might correct for the different levels of inflammatory cytokines between patients and their possible influence on the expression of CD4+ T-cells. In this respect, it is notable that CRP was already significantly different between ETN responders and non-responders in the microarray cohort. Correction for baseline DAS28 might correct for gene expression levels and the (inseparable) a priori chances on achieving a EULAR response. Although there is no consensus over the confounders yet, prediction models should consider including the most influencing parameters, such as CRP and DAS28, to increase generalizability.

## **CONCLUSION**

Our study did not identify gene transcripts consistently predictive for response to either ADA or ETN across different cohorts. However, two genes with apoptosis related functions, IQSEC1 and BTF3, hold some potential as they predicted response in two of three ETN cohorts. The presence of technical variability in microarray observed in this study and previous research, may play a role in the difficulty to reproduce results. Since RNA-seq showed less variability, the choice for this technique might constitute a better option for future studies. The large clinical heterogeneity in RA might additionally explain some of the differences in results, and correction for confounding factors may improve the generalizability of future results further.

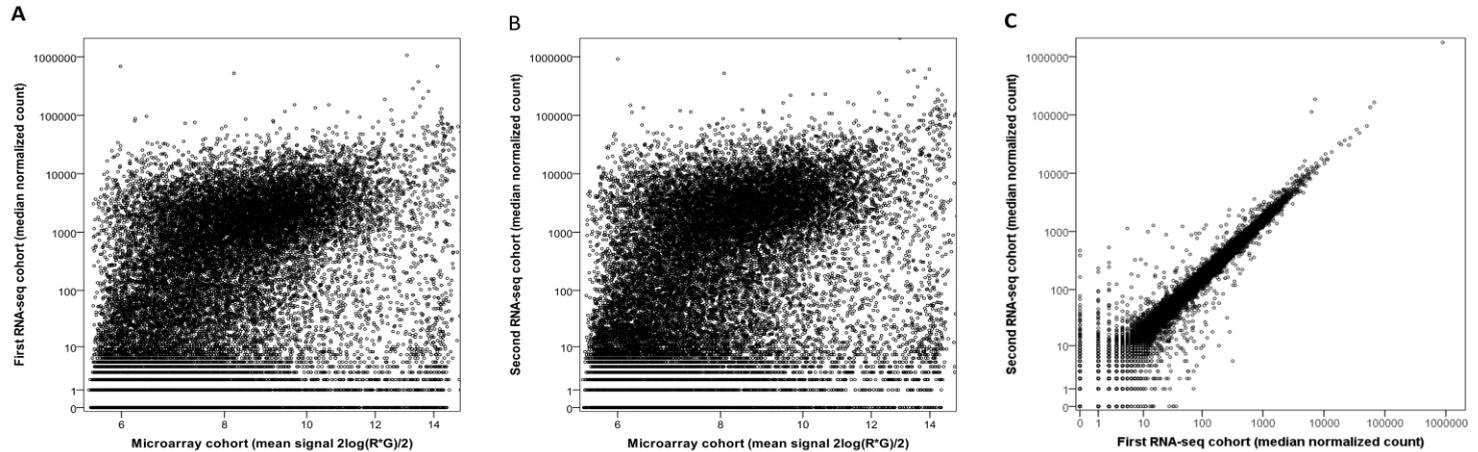
## REFERENCES

1. Keystone EC, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum* 2004;50 5: 1400-11.
2. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med* 1999;340 4: 253-9.
3. Hatfield GW, Hung SP, Baldi P. Differential analysis of DNA microarray gene expression data. *Mol Microbiol* 2003;47 4: 871-7.
4. Badot V, Galant C, Nzeusseu TA, Theate I, Maudoux AL, Van den Eynde BJ et al. Gene expression profiling in the synovium identifies a predictive signature of absence of response to adalimumab therapy in rheumatoid arthritis. *Arthritis Res Ther* 2009;11 2: R57.
5. Bansard C, Lequerre T, Derambure C, Vittecoq O, Hiron M, Daragon A et al. Gene profiling predicts rheumatoid arthritis responsiveness to IL-1Ra (anakinra). *Rheumatology (Oxford)* 2011;50 2: 283-92.
6. Dennis G, Jr., Holweg CT, Kummerfeld SK, Choy DF, Setiadi AF, Hackney JA et al. Synovial phenotypes in rheumatoid arthritis correlate with response to biologic therapeutics. *Arthritis Res Ther* 2014;16 2: R90.
7. Julia A, Erra A, Palacio C, Tomas C, Sans X, Barcelo P et al. An eight-gene blood expression profile predicts the response to infliximab in rheumatoid arthritis. *PLoS One* 2009;4 10: e7556.
8. Julia A, Barcelo M, Erra A, Palacio C, Marsal S. Identification of candidate genes for rituximab response in rheumatoid arthritis patients by microarray expression profiling in blood cells. *Pharmacogenomics* 2009;10 10: 1697-708.
9. Kim TH, Choi SJ, Lee YH, Song GG, Ji JD. Gene expression profile predicting the response to anti-TNF treatment in patients with rheumatoid arthritis; analysis of GEO datasets. *Joint Bone Spine* 2014;81 4: 325-30.
10. Koczan D, Drynda S, Hecker M, Drynda A, Guthke R, Kekow J et al. Molecular discrimination of responders and nonresponders to anti-TNF alpha therapy in rheumatoid arthritis by etanercept. *Arthritis Res Ther* 2008;10 3: R50.
11. Lee YH, Bae SC, Song GG. Meta-analysis of gene expression profiles to predict response to biologic agents in rheumatoid arthritis. *Clin Rheumatol* 2014;33 6: 775-82.
12. Lequerre T, Gauthier-Jauneau AC, Bansard C, Derambure C, Hiron M, Vittecoq O et al. Gene profiling in white blood cells predicts infliximab responsiveness in rheumatoid arthritis. *Arthritis Res Ther* 2006;8 4: R105.
13. Lindberg J, af KE, Catrina AI, Nilsson P, Klareskog L, Ulfgren AK et al. Effect of infliximab on mRNA expression profiles in synovial tissue of rheumatoid arthritis patients. *Arthritis Res Ther* 2006;8 6: R179.
14. Lindberg J, Wijbrandts CA, van Baarsen LG, Nader G, Klareskog L, Catrina A et al. The gene expression profile in the synovium as a predictor of the clinical response to infliximab treatment in rheumatoid arthritis. *PLoS One* 2010;5 6: e11310.
15. MacIsaac KD, Baumgartner R, Kang J, Loboda A, Peterfy C, DiCarlo J et al. Pre-treatment whole blood gene expression is associated with 14-week response assessed by dynamic contrast enhanced magnetic resonance imaging in infliximab-treated rheumatoid arthritis patients. *PLoS One* 2014;9 12: e113937.
16. Meugnier E, Coury F, Tebib J, Ferraro-Peyret C, Rome S, Bienvenu J et al. Gene expression profiling in peripheral blood cells of patients with rheumatoid arthritis in response to anti-TNF-alpha treatments. *Physiol Genomics* 2011;43 7: 365-71.
17. Nakamura S, Suzuki K, Iijima H, Hata Y, Lim CR, Ishizawa Y et al. Identification of baseline gene expression signatures predicting therapeutic responses to three biologic agents in rheumatoid arthritis: a retrospective observational study. *Arthritis Res Ther* 2016;18: 159.
18. Oliveira RD, Fontana V, Junta CM, Marques MM, Macedo C, Rassi DM et al. Differential gene expression profiles may differentiate responder and nonresponder patients with rheumatoid arthritis for methotrexate (MTX) monotherapy and MTX plus tumor necrosis factor inhibitor combined therapy. *J Rheumatol* 2012;39 8: 1524-32.
19. Oswald M, Curran ME, Lamberth SL, Townsend RM, Hamilton JD, Chernoff DN et al.

- Modular analysis of peripheral blood gene expression in rheumatoid arthritis captures reproducible gene expression changes in tumor necrosis factor responders. *Arthritis Rheumatol* 2015;67 2: 344-51.
20. Raterman HG, Vosslander S, de RS, Nurmohamed MT, Lems WF, Boers M et al. The interferon type I signature towards prediction of non-response to rituximab in rheumatoid arthritis patients. *Arthritis Res Ther* 2012;14 2: R95.
  21. Sekiguchi N, Kawauchi S, Furuya T, Inaba N, Matsuda K, Ando S et al. Messenger ribonucleic acid expression profile in peripheral blood cells from RA patients following treatment with an anti-TNF-alpha monoclonal antibody, infliximab. *Rheumatology (Oxford)* 2008;47 6: 780-8.
  22. Stuhlmüller B, Haupl T, Hernandez MM, Grutzkau A, Kuban RJ, Tandon N et al. CD11c as a transcriptional biomarker to predict response to anti-TNF monotherapy with adalimumab in patients with rheumatoid arthritis. *Clin Pharmacol Ther* 2010;87 3: 311-21.
  23. Tanino M, Matoba R, Nakamura S, Kameda H, Amano K, Okayama T et al. Prediction of efficacy of anti-TNF biologic agent, infliximab, for rheumatoid arthritis patients using a comprehensive transcriptome analysis of white blood cells. *Biochem Biophys Res Commun* 2009;387 2: 261-5.
  24. Thomson TM, Lescarbeau RM, Drubin DA, Laifenfeld D, de GD, Fryburg DA et al. Blood-based identification of non-responders to anti-TNF therapy in rheumatoid arthritis. *BMC Med Genomics* 2015;8: 26.
  25. Toonen EJ, Gilissen C, Franke B, Kievit W, Eijssbouts AM, den Broeder AA et al. Validation study of existing gene expression signatures for anti-TNF treatment in patients with rheumatoid arthritis. *PLoS One* 2012;7 3: e33199.
  26. van der Pouw Kraan TCM, Wijbrandts CA, van Baarsen LG, Rustenburg F, Baggen JM, Verweij CL et al. Responsiveness to anti-tumour necrosis factor (alpha) therapy is related to pre-treatment tissue inflammation levels in rheumatoid arthritis patients. *Ann Rheum Dis* 2008;67 4.
  27. Vosslander S, Raterman HG, TC vdPK, Schreurs MW, von Blomberg BM, Nurmohamed MT et al. Pharmacological induction of interferon type I activity following treatment with rituximab determines clinical response in rheumatoid arthritis. *Ann Rheum Dis* 2011;70 6: 1153-9.
  28. Fu X, Fu N, Guo S, Yan Z, Xu Y, Hu H et al. Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics* 2009;10: 161.
  29. 't Hoen PA, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RH, de Menezes RX et al. Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Res* 2008;36 21: e141.
  30. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* 2008;18 9: 1509-17.
  31. Wright HL, Thomas HB, Moots RJ, Edwards SW. Interferon gene expression signature in rheumatoid arthritis neutrophils correlates with a good response to TNFi therapy. *Rheumatology (Oxford)* 2015;54 1: 188-93.
  32. Folkersen L, Brynedal B, Diaz-Gallo LM, Ramskold D, Shchetynsky K, Westerlind H et al. Integration of known DNA, RNA and protein biomarkers provides prediction of anti-TNF response in rheumatoid arthritis: results from the COMBINE study. *Mol Med* 2016;22.
  33. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ* 2009;338: b375.
  34. Bellucci E, Terenzi R, La Paglia GM, Gentileschi S, Tripoli A, Tani C et al. One year in review 2016: pathogenesis of rheumatoid arthritis. *Clin Exp Rheumatol* 2016;34 5: 793-801.
  35. van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996;39 1.
  36. van Wageningen S, Kemmeren P, Lijnzaad P, Margaritis T, Benschop JJ, de Castro IJ et al. Functional overlap and regulatory links shape genetic interactions between signaling pathways. *Cell* 2010;143 6: 991-1004.
  37. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 2002;30 4: e15.
  38. Wu H, Kerr MK, Cui X, Churchill GA. MAANOVA: A software package for the

- analysis of spotted cDNA microarray experiments. 2003:  
<http://research.jax.org/faculty/churchill>.
39. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010;11 10: R106.
  40. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* 2014;15 2: R29.
  41. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43 7: e47.
  42. Oliveros JC. Venny. An interactive tool for comparing lists with Venn's diagrams. 2007-2015:  
<http://bioinfo.cnb.csic.es/tools/venny/index.html>.
  43. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 2009;37 Web Server issue: W305-11.
  44. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet* 2016.
  45. Smolen JS, Aletaha D. Forget personalised medicine and focus on abating disease activity. *Ann Rheum Dis* 2013;72 1: 3-6.
  46. Someya A, Sata M, Takeda K, Pacheco-Rodriguez G, Ferrans VJ, Moss J et al. ARF-GEP(100), a guanine nucleotide-exchange protein for ADP-ribosylation factor 6. *Proc Natl Acad Sci U S A* 2001;98 5: 2413-8.
  47. Zhu W, London NR, Gibson CC, Davis CT, Tong Z, Sorensen LK et al. Interleukin receptor activates a MYD88-ARNO-ARF6 cascade to disrupt vascular stability. *Nature* 2012;492 7428: 252-5.
  48. Ouyang S, Hsueh H, Kastin AJ, Pan W. TNF stimulates nuclear export and secretion of IL-15 by acting on CRM1 and ARF6. *PLoS One* 2013;8 8: e69356.
  49. Someya A, Moss J, Nagaoka I. Involvement of a guanine nucleotide-exchange protein, ARF-GEP100/BRAG2a, in the apoptotic cell death of monocytic phagocytes. *J Leukoc Biol* 2006;80 4: 915-21.
  50. Liu Q, Zhou JP, Li B, Huang ZC, Dong HY, Li GY et al. Basic transcription factor 3 is involved in gastric cancer development and progression. *World J Gastroenterol* 2013;19 28: 4495-503.
  51. Wang CJ, Franbergh-Karlson H, Wang DW, Arbman G, Zhang H, Sun XF. Clinicopathological significance of BTF3 expression in colorectal cancer. *Tumour Biol* 2013;34 4: 2141-6.
  52. Symes AJ, Eilertsen M, Millar M, Nariculam J, Freeman A, Notara M et al. Quantitative analysis of BTF3, HINT1, NDRG1 and ODC1 protein over-expression in human prostate cancer tissue. *PLoS One* 2013;8 12: e84295.
  53. Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73 3: 492-509.
  54. Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol* 2016;68 1: 1-26.
  55. Shabihkhani M, Lucey GM, Wei B, Mareninov S, Lou JJ, Vinters HV et al. The procurement, storage, and quality assurance of frozen blood and tissue biospecimens in pathology, biorepository, and biobank settings. *Clin Biochem* 2014;47 4-5: 258-66.
  56. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 2009;338: b606.
  57. Steyerberg EW. *Clinical Prediction Models*. Publisher: Springer Science 2009: ISBN: 978-0-387-77243-1.

**Supplementary Figure 1:** Correlation between read counts of all expressed genes. In order to test the test-retest reliability, the correlations between gene expression as measured by microarray, first and second RNA seq study were compared for overlapping patients. Per gene transcript, the expression was plotted on log scales, and the correlation was calculated. (A) microarray *versus* RNA-seq pilot, n=14, r=0.401, p<0.0001 (B) microarray versus RNA-seq full study, n=13, r = 0.397, p<0.0001 (C): RNA seq pilot versus RNA-seq full study, n=5, r=0.935, p<0.0001.







## Chapter 7

### **Proteomics to predict the response to TNF-alpha inhibitors in rheumatoid arthritis using a supervised cluster-analysis based protein score**

Bart V.J. Cuppen, Ruth D.E. Fritsch-Stork, Iris Eekhout, Wilco de Jager, Anne C. Marijnissen, Johannes W.J. Bijlsma, Martijn Custers, Jacob M. van Laar, Floris P.J.G. Lafeber, and Paco M.J. Welsing, on behalf of all SRU-investigators

[Minor revision]

## **ABSTRACT**

### **Objective**

In rheumatoid arthritis (RA) it is of major importance to distinguish non-responders to TNF-alpha inhibitor (TNFi) treatment before start to prevent a delay in effective treatment. We developed a protein score for the response to TNFi treatment in RA and investigated its predictive value.

### **Methods**

In RA-patients eligible for biological treatment were included in the BiOCURA registry, 53 inflammatory proteins were measured using xMAP technology. A supervised cluster analysis method, partial least squares (PLS), was used to select the best combination of proteins. Using logistic regression, a predictive model containing readily available clinical parameters was developed and the potential of the this model with and without the protein score to predict EULAR response was assessed using the area under the receiving operator curve (AUC-ROC) and the net reclassification index (NRI).

### **Results**

For the development step (n=65 patient), PLS revealed twelve important proteins: CCL3 (MIP1a), CCL17 (TARC), CCL19 (MIP3b), CCL22 (MDC), IL4, IL6, IL7, IL15, sCD14, sCD74 (MIF), sIL1-RI, and sTNF-RII. The protein score scarcely improved the AUC-ROC (0.72 to 0.77), and the ability to improve classification and reclassification (NRI=0.05). In validation (n=185), the model including protein score did not improve the AUC-ROC (0.71 to 0.67), or the reclassification (NRI=-0.11).

### **Conclusion**

No proteomic predictors were identified that were better suitable than clinical parameters to distinguish TNFi non-responders from responders before the start of treatment. As the results between previous studies and this study are disparate, we currently have no proteomic predictors for the response to TNFi.

## **INTRODUCTION**

Rheumatoid arthritis (RA) is a chronic, disabling disease that mainly affects the synovial joints. The disease is autoimmune in origin and affects 0.5-1% of the population in industrialized countries[1, 2]. Although the introduction of tumor necrosis factor- $\alpha$ -inhibiting therapy (TNFi) has dramatically improved the outlook for RA patients a substantial proportion of patients (approximately 30-40%) fail to respond to TNFi therapy[3, 4]. As we currently cannot differentiate at the start of therapy between patients that will respond and patients that will not [5], TNFi treatment is administered in an empirical trial and error approach. This entails therapeutic inefficiency in a substantial part of patients, due to the three to six months observation period before clinical response can be judged. In this time frame, the non-responding patient may suffer from uncontrolled disease with the potential of irreversible damage and harmful side effects from treatment. One other problem is the high costs incurred with this treatment. The challenge is therefore to identify responders and non-responders to specific therapies beforehand, and to initiate optimal treatment.

TNFi targets TNF- $\alpha$ , which is central in the pathogenesis of RA and acts within a complex network of cells and cytokines that relate to inflammation[6]. It is therefore likely that proteomic approaches, in which many cytokines are measured simultaneously and are jointly investigated, are capable of classifying RA patients at baseline according to their response to TNFi treatment. We explored the baseline serum proteome on predictors that might identify responders and non-responders to TNFi, using a panel of proteins that were all specifically related to inflammation, and aimed to validate our models in a new cohort of consecutively included patients. We investigated whether a combination of proteins and clinical parameters predict response to TNFi treatment in RA patients better than clinical parameters alone.

## **METHODS**

### **Patients**

Patients that were initiated on TNFi treatment were selected from the BiOCURA cohort (Biologicals and Outcome, Compared and predicted in Utrecht region, in Rheumatoid

Arthritis). BiOCURA is an observational cohort, in which RA patients eligible for biological treatment according to regular clinical practice were enrolled and followed up until one year after start of treatment, in one academic hospital and seven regional hospitals in the Netherlands. In order to reflect clinical practice, no patients were excluded beforehand. Re-inclusion after switching to a different biological treatment was possible. The study was approved by the local ethics committee of the University Medical Center Utrecht and the institutional review boards of the participating centers (see Acknowledgements). Informed consent was obtained from each patient.

### **Measurements**

Trained nurses gathered all data during a separate visit, which included all clinical parameters, joint counts and blood, before the first dose of the treatment was given. Blood was processed immediately after collection and serum samples were stored at -80 degrees Celsius until analysis. Serum levels of 53 inflammatory proteins were measured using xMAP technology (Luminex, Texas, USA), a multiplex bead-based immunoassay that allows the measurement of a large number of proteins simultaneously. Before xMAP analysis was performed, non-specific heterophilic antibodies, such as rheumatoid factors, were pre-absorbed from all samples with protein-L pre-coated enzyme-linked immunosorbant assay (ELISA) plates[7] to minimize interference with the immunoassay. The selection of 53 proteins, is displayed in **Supplementary Table 1**. The xMAP analyses for the development and validation were performed in respectively October 2014 and June 2016.

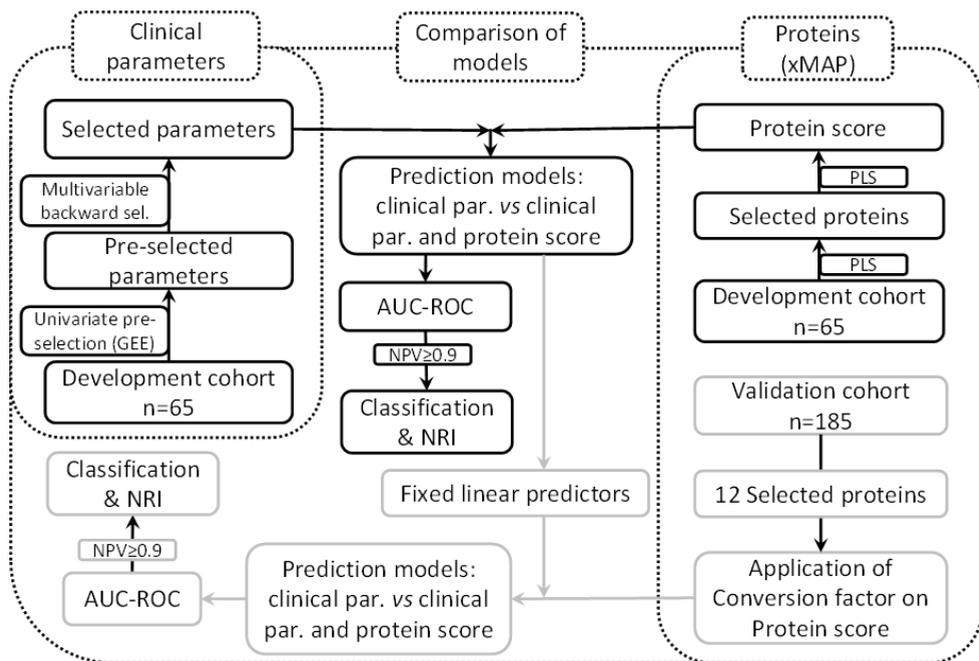
### **Outcome**

Disease activity was assessed using the disease activity score based on a 28-joint count (DAS28)[8]. Response to treatment was based on the European League Against Rheumatism (EULAR) response criteria[9]. The EULAR distinguishes a non-, moderate- and good response. As our main focus was to predict a good-response in contrast to a moderate response, non- and moderate-response were combined and compared with good-response. EULAR good response is defined as a DAS28 improvement  $\geq 1.2$  and an absolute DAS28  $\leq 3.2$  after treatment. Patients that discontinued therapy before three

months due to side effect, were excluded from the study because they cannot be considered merely missing at random and are not necessarily non-responders.

### **Analysis**

All treatment courses with TNFi treatment (i.e. adalimumab, etanercept, golimumab, infliximab or certolizumab) were included for this study, including courses from patients having switched from one to another TNFi therapy within BiOCURA. 276 TNFi treatment courses were divided into separate groups based on the treatment starting dates. The development of the models was performed on cases included between October 2012 and March 2014 (n=65), whereas validation was performed on TNFi treatment cases included from June 2009 and October 2012 (n=185). Development and validation were performed on cohorts split by treatment onset dates (i.e. non-random split), which leads to a different clinical setting for the inclusion of patients per cohort. Therefore, the validation can be considered in-between internal and external validation[10]. Details for missing data handling by multiple imputations can be found in **Supplementary Text 1**. The statistical analyses of the development, refinement and validation step are described in the subsequent paragraphs and summarized in **Figure 1**.



**Figure 1:** Overview of statistical analyses. The pathways denoted by black-dotted and grey-straight lines represent the analyses performed in the development and validation step respectively. AUC-ROC, receiving operator characteristic curve; GEE, generalized estimation equations; NPV, negative predictive value; NRI, net reclassification index; PLS, partial least squares

### Development

In the development step, proteins were measured in 65 samples and log-transformed to approximate a normal distribution. These log-transformed values were used in a partial least squares (PLS) analysis, a supervised clustering algorithm, to select the most contributing proteins and build the protein score that predicts DAS28 at three months. Those proteins not normally distributed were not entered in PLS as they not only disturb PLS, but are also unfeasible regarding (regression based) risk prediction. The inflammation marker C-reactive protein (CRP, log-transformed value) and DAS28 at baseline were always included, because they could be influenced by previous treatment and in turn influence the inflammatory profile and classification of EULAR response. The most important proteins were selected based on variable importance in projection (VIP)-scores  $\geq 1$  in combination with the highest factor loadings, so that no more than 15 proteins were

included; Subsequently, this selection of proteins together with baseline CRP and DAS28 were included in the PLS procedure again to create the initial protein score. In this procedure, the PLS algorithm suggests the best linear fit of the proteins with the DAS28 at three months, considering the results of the supervised clustering, e.g.  $1.5 * \text{Ln}(\text{TNF-alpha}) + 0.7 * \text{Ln}(\text{IL-6})$  etc.

The inclusion of patients that were re-included after switching to a different TNFi resulted in two-level data (i.e. data from different patient as well as repeated data within patients), which we accounted for by using generalized estimating equation (GEE) modelling in the pre-selection of clinical parameters[11], This univariate analysis was carried out for all clinical baseline variables (Table 1) including the health-related quality of life questionnaire (EQ5D) and health assessment questionnaire (HAQ), using GEE with logit link function, and parameters were selected with a p-value <0.2. All selected clinical parameters were then entered in a (full) logistic regression model and were removed one-by-one using a backward selection procedure ( $p < 0.1$ ), while baseline CRP and DAS28 were always included, to form the “clinical” model. These clinical parameters were also used as covariates together with the protein score in the “combined” model. Comparison of the clinical and combined model allows determination of the added value of measuring proteins for prediction of response to TNFi additional to parameters already available in clinical practice. The discriminating potential of the models was assessed using the area under the receiving operator characteristic curve (AUC-ROC). Then per model a cut-off was chosen based on a negative predictive value (NPV) higher than 90%. If patients fail to meet this cut-off their probability of good response is <10%, which was considered a clinically relevant cut-off to consider an alternative (combination) treatment. All these patients were classified to the ‘low probability of response’ category and all patients that did meet this cut-off were assigned to the ‘normal probability of response’ category. Based on these categories, the number of patients that were classified to this low probability of response category was compared between the models, to determine whether the combined model is more sensitive to classify non-responders. As an indication, a number needed to test (NNT) by the combined model before one more patient is classified into the low category of response was calculated. Additionally and also based on these categories, the net reclassification index (NRI) was calculated to determine

the ability of the combined model to reclassify patients as predicted by the clinical model into a (more) correct category. For example, when a future non-responder is classified at baseline to have a normal probability of response by the model with clinical model, but is classified into the low probability category in the combined model, the reclassification is “more correct” and the NRI for non-responders increases, as well as the total NRI (=NRI for responders + NRI non-responders).

### Validation

In the validation step, the selected proteins from the development step were measured with xMAP technology in an additional 185 patients. Subsequently, the two prediction models from the development step were recalibrated by entering the prediction rule (i.e. the linear predictor from the models developed, e.g.  $0.8 * \text{protein score} + 2 * \text{age} + 3 * \text{gender}$  etc.) in a logistic regression model.

The AUC-ROC was calculated and compared between the two models (i.e. the clinical and the combined model) and after setting the pre-defined cut-offs the added value of the protein score was analyzed by comparing the classification and the NRI.

### Used software

The PLS analyses were performed in SAS version 9.2 (SAS institute Inc. Cary NC, USA), and the multiple imputation as well as the development and validation steps of the prediction models in SPSS version 21.0.0.0. (SPSS Inc., Chicago, Illinois, USA).

## **RESULTS**

### **Baseline characteristics**

Baseline characteristics for patients of the development, refinement and validation cohorts are shown in Table 1. Notable differences in the validation cohort, which was a few years earlier in time compared to the development cohort, were a higher CRP (median 6.0 *versus* 4.0,  $p=0.011$ ), higher DAS28 (4.5 *versus* 4.1,  $p=0.019$ ) and lower concomitant MTX use (71.9 *versus* 98.5%,  $p<0.001$ ). Preferences for certain TNFi

treatments changed over time due to commercial influences, e.g. more etanercept was administered in recent years (i.e. development cohort).

**Table 1:** Baseline characteristics of development (n=65) and validation cohort (n=185) before multiple imputation.

Clinical parameter	Development (n=65)	Validation (n=185)	p-value
Gender, female, n (%)	47 (72.3)	138 (74.6)	0.744
Age, mean (SD)	57.0 (±10.9)	54.1 (±12.3)	0.094
Disease duration, years, median (IQR)	6.0 (2.0-14.5)	6.0 (2.0-12.0)	0.817
Smoking currently, n (%)	23 (35.4)	44 (23.8)	0.075
RF, positive, n (%)	36 (55.4)	122 (66.7)	0.133
ACPA, positive, n (%)	45 (69.2)	131 (72.4)	0.634
CRP, mg/L, median (IQR)	4.0 (1.1-11.0)	6.0 (3.0-13.0)	<b>0.011</b>
ESR, mm/hr, median (IQR)	14.0 (7.0-33.5)	19.5 (10.0-33.8)	0.145
Baseline DAS28, mean (SD)	4.1 (±1.2)	4.5 (±1.2)	<b>0.019</b>
No. of previously used bDMARDs, n (%)			0.520
0 (naïve)	35 (53.8)	118 (63.8)	
1	22 (33.8)	53 (28.6)	
≥2	8 (12.3)	14 (7.6)	
Initiating treatment with, n (%)			0.058
Adalimumab	17 (26.2)	73 (39.5)	
Etanercept	33 (50.8)	65 (35.1)	
Golimumab	5 (7.7)	26 (14.1)	
Infliximab	3 (4.6)	11 (5.9)	
Certolizumab	7 (10.8)	10 (5.4)	
Concomitant treatment, n (%)			
MTX	64 (98.5)	133 (71.9)	<b>&lt;0.001</b>
GCs	24 (36.9)	67 (36.2)	1.000

ACPA, anti-citrullinated protein antibody; bDMARD, biological DMARD; DMARDs, disease modifying anti-rheumatic drugs; csDMARD, conventional synthetical DMARD (incl. GCs); IQR, inter-quartile range; GCs, glucocorticoids; MTX, methotrexate; RF, rheumatoid factor; SD, standard deviation.

### Development of protein score

Of all 53 proteins, 38 were near normally distributed after logtransformation and were used for PLS analysis. Regarding the other 15 proteins that were rejected, usually many values fell below the detection limit, hampering the usage of these protein measurements in risk prediction. By selection on  $VIP \geq 1$  and a factor loading of  $>0.135$ , twelve proteins were selected: CCL3 (macrophage inflammatory protein 1a, MIP1a), CCL17 (thymus and

activation regulated chemokine, TARC), CCL19 (macrophage inflammatory protein MIP3b), CCL22 (macrophage-derived chemokine, MDC), interleukin 4 (IL4), IL6, IL7, IL15, soluble cluster of differentiation 14 (sCD14), sCD74 (macrophage migration inhibitory factor, MIF), soluble IL1 receptor I (sIL1-RI), and soluble tumor necrosis factor receptor II (sTNF-RII). A protein score was created by entering these proteins in PLS again (Supplementary Table 2).

### **Development of models**

After univariate and multivariable selection, the clinical parameters CRP, DAS28 and the health assessment questionnaire (HAQ) remained and a prediction model with and without the protein score was fitted to the data (see Table 2). The clinical model predicted response with an AUC-ROC of 0.72 (0.64 – 0.80), whereas the combined model including the protein score resulted in an AUC-ROC of 0.77 (0.70 – 0.85). When the prediction models were applied to classify patients into response categories, the combined model was able to assign more patients into the category of low probability on EULAR good response with 28.2% *versus* 21.5% respectively, with a NNT of 15 before one extra patient was classified into the low category (see Table 3). When the redistribution of individual patients was investigated after addition of the protein score, the reclassification of non-responders was 8.2% better and of responders 3.3% lower, resulting in a NRI of 0.05.

**Table 2:** Models for (baseline) prediction of response in development cohort (n=65). The CRP was logtransformed to approximate a normal distribution. The clinical model, containing only clinical parameters, showed an AUC-ROC (95%-CI) of 0.72 (0.59 – 0.86). The combined model, containing clinical parameters and the protein score, showed an AUC-ROC (95%-CI) of 0.77 (0.64 – 0.90)

Item	Clinical model		Combined model	
	OR (95%-CI)	P-value	OR (95%-CI)	P-value
<b>Protein Score</b>			0.24 (0.07-0.83)	0.024
<b>DAS28 at baseline</b>	1.64 (0.89-3.03)	0.112	3.18 (1.30-7.77)	0.011
<b>Ln (CRP)</b>	1.32 (0.76-2.28)	0.321	1.04 (0.56-1.93)	0.894
<b>HAQ</b>	0.30 (0.08-1.13)	0.075	0.33 (0.08-1.43)	0.138

AUC-ROC, area under the curve of the receiving operator curve; CI, confidence interval; CRP, C-reactive protein; DAS28, disease activity score based on 28 joints; HAQ, health assessment questionnaire; OR, odds ratio.

**Table 3:** Categories for probability on EULAR good response and estimated group size per model in development cohort (n=65). The low category are based on cut-offs for NPV>0.9 (resulting in a probability on good response <10% for the low category). The number needed to test (NNT) represents the number of patients needed to be tested with proteins to put one extra patient in the low probability of response category.

Category for probability of good response	Clinical model	Combined model
<b>Low (&lt;10%), estimated size %</b>	21.5	28.2
<b>Normal (&gt;10%), estimated size %</b>	79.5	71.8
<b>NNT</b>		15

### Validation of models

We noticed on the xMAP analyses of the validation cohort, that there was a decay in concentration for 10/12 proteins (especially the interleukins), which could well relate to the longer storage time at -80°C (see Supplementary Table 3). In order not to involuntarily adjust the relative weight of each protein in the protein score, the protein levels in the validation cohort had to be adjusted so that the mean and standard deviations were comparable with the development step. We did so by converting the log transformed value in the validation cohort to a Z-score with mean=0 and SD=1, and subsequently multiply these Z-score with the SD of the corresponding protein in the development cohort and add the mean value of the protein in the development cohort. These adjusted proteins were then put in the original protein-score, without any refitting of the relative

weight between proteins or refitting with respect to the outcome. The calculated prediction score from the clinical and combined model were applied in the validation cohort, which resulted in an AUC-ROC of 0.71 (0.63 – 0.79) and 0.67 (0.59 – 0.75) for the model without- and with the protein score, respectively. The model including the protein score was not able to classify more patients into the low (17.3% versus 6.1% of total) probability of response category (see Table 4) and the NNT of -9 indicates that for every nine patients classified by the combined model, one less patient is classified into the low probability of response category. When the reclassification was investigated, classification of non-responders was 14.6% lower and of responders 3.5% higher, resulting in a NRI of -0.11.

**Table 4:** Risk categories for EULAR good response and estimated group size per model in validation cohort (n=185). The low category are based on cut-offs for NPV>0.9 (resulting in a probability on good response <10% for the low category). The number needed to test (NNT) represents the number of patients needed to be tested with proteins to put one extra patient in the low probability of response category.

Category for probability of good response	Clinical model	Combined model
Low (<10%), estimated size %	17.3	6.1
Normal (>10%), estimated size %	82.7	93.9
NNT		-9

## DISCUSSION

In this study we aimed to develop a protein score that is predictive for the response to TNFi treatment in RA, and to investigate whether this score is of added predictive value over using clinical parameters alone. We showed that the protein score combined with clinical parameters initially improved the prediction of the TNFi treatment response compared with clinical parameters alone (including DAS28, CRP and the HAQ), though this could not be validated in a separate cohort of consecutively included patients. In fact, the clinical model even outperformed the combined model in the validation cohort. Using DAS28, CRP and HAQ, which are parameters easily accessible in the clinical setting, 17% of the validation cohort could be classified to the low risk category (<10%). However, it is

questionable if these are truly specific parameters, as the low DAS28 already limits options for a good response considerably (more difficult to achieve a DAS28 change >1.2) and a high HAQ is to some extent indicative for treatment refractory RA. It is therefore not expected that the response of patients with a low DAS28 and high HAQ to any non-TNFi will be better.

Proteomic approaches for prediction of response have been performed before [12-19]. An overview of these studies and reported results is presented in Table 5. Two studies found no predictors for response, of which one study did not identify any possible predictors in the discovery phase [14], whereas the other study failed to validate the results of the discovery phase [18]. The other 6 studies all proposed predictors for response to therapy, however, none of these studies performed a technical replication or separate validation in new patients, and the overlap of discovered predictors between studies is restricted to CCL2 (monocyte chemoattractant protein 1, MCP-1), which has been cross validated in two studies[12, 13]. However, CCL2 did not show to be predictive in our study and was not identified in the studies employing mass spectrometry [15, 16], implying that it is not a consistent predictor. Several other proteins proposed as predictors in earlier studies were measured in our xMAP panel, including CXCL4, GM-CSF, IL-1 $\alpha$ , IL-1 $\beta$ , IL-6, IL-8, IL-15, IL-19, IP-10 and TNF $\alpha$ . However, none of these parameters were associated with response in the development step, except for IL-15 which could not be validated and was also not identified in the studies using mass spectrometry. The disparate results between all studies to date may be explained by differences in demographic characteristics, serology status, used TNFi and/or concomitant treatments and different techniques, however the possibility of false-positive findings and absence of biologically true predictors for response cannot be excluded.

**Table 5:** Non-exhaustive overview of studies with a proteomic approach to identify baseline predictive proteins regarding response to TNFi treatment. Proteins in bold indicates this protein is overlapping with another study.

Study (et al.)	Treatment	No. of patients	Material	Method	No. of proteins	Found associations	Tech repl*	Validation in new patients
<b>Fabre [12]</b>	ETN	33	Serum	Protein Biochip Array Technology (Investigator Evidence, Randox, Manguio, France)	12	High EGF & <b>MCP-1</b> predicts good response	No	No
<b>Hueber [13]</b>	ETN	93	Serum	12-plex human cytokine FLEX® kit (Millipore) and ELISA (targets pre-selected by autoantigen microarrays)	26	24 protein signature predicts response (incl. GM-CSF-, IL6, IL1a, IL1b, Eotaxin, IP-10, FGF-2, <b>MCP-1</b> , IL12p40, IL12p70, IL15, Fibrinogen protein, FibA (616-635)cit, Fibromodulin, Clusterin (2x), ApoE, H2B/e, HSP58, COMP, Acetyl-calpastatin, Biglycin, Osteoglycin, SerineProtease-11, unknown direction)	No	No
<b>Lequerre [14]</b>	IFX	76	Serum	ELISA, high performance liquid chromatography, latex agglutination and Waaler-Rose tests	22	No predictors	No	No
<b>Ortea [15]</b>	IFX	8	Serum	LC-MS/MS	287	Low concentrations of CP, GC, apoA-II, apoB-100, apoM, ITIH1, ITIH2, GSN, CFHR4, APMAP, MASP2, thrombospondin-1 and complement C4-B alpha chain predict response	No	No
<b>Trocme [16]</b>	IFX	60	Plasma	SELDI-TOF MS	NA (infinite)	High concentration of six biomarkers predict response, of which two proteins identified: apoA-I and CXCL4 (PF4)	No	No

<b>Uno [17]</b>	ETN	43	Serum	Bio-Plex Human Cytokine 27-Plex panel (BioRad) and Milliplex MAP (EDM Millipore)	31	High IL9 and low TNFa predict high DAS28-CRP	No	No
<b>Visvanathan [18]</b>	GLM	100	Serum	Luminex, ELISA and HumanMAP (Myriad RBM)	107	Low concentrations of vWF and pyridinoline predict ACR20 response	No	Results could not be validated in 236 new patients
<b>Wright [19]</b>	IFX, ADA, ETN	22	Synovial fluid	human cytokine Luminex multiplex assay (Biosource, Paisley, UK)	12	High IL2 and G-CSF and low IL6 predict response	No	No

\*technical replication of protein measurements after final selection / developed model

The results should be viewed within the perspective of the limitations of this study. First, we used GEE to adjust for the extra level in the data, however we were unable to use this technique in validation as no universal prediction score (i.e. linear predictor) can be derived from the GEE model that can be applied in another cohort. As a sensitivity analysis, therefore, we fitted a clinical and combined model using GEE to the data in the validation cohort, containing the individual parameters instead of the prediction scores. The AUC-ROC of both models was 0.72 (0.64-0.80), again indicating that there is no advantage of the inclusion of the protein score. Another weak point of this study is the large differences in protein concentration between the development and validation cohort which was probably related to a longer storage time and subsequent degradation. The means and SDs of the (logtransformed) protein concentrations were therefore rescaled to the values of the newest samples in order not to influence the protein score. However, one important assumption for this adjustment is that the actual protein concentrations of the patients in the validation cohort are likely to follow the same distribution as in the development cohort, which might not be the case considering the higher baseline CRP and DAS28 in the validation cohort. However, as none of the selected proteins significantly correlated with baseline DAS28 (all  $p < 0.05$ ), and only IL6 ( $r = 0.255$ ,  $p = 0.042$ ) and sCD14 ( $r = 0.294$   $p = 0.018$ ) weakly correlated with CRP, the expected influence of baseline characteristics in the observed differences in protein concentrations between the development and validation cohort is small.

Although proteomic approaches thus far fail to provide predictors of response to treatment, they are capable of assessing current disease activity. Moreover, a multi-biomarker disease activity (MBDA) test for rheumatoid arthritis was developed [20] and validated in several external cohorts [21-24]. Although careful clinical judgement cannot be replaced by any biomarker assay, the MBDA test can possibly be used to biochemically track changes in disease activity over time. In early RA, the baseline MBDA was a significant prognostic factor for radiographic progression in one study [24], yet not in a different study [21]. In our study we were not able to test its predictive ability for therapeutic response, as the individual parameters within the MBDA were not all covered by the proteomic platform.

Heterogeneity plays a large role in RA research. Patients differ in all aspects of the disease, e.g. genetic susceptibility, autoantibody status, erosive disease, pain perception, previous failed treatments, efficacy and tolerability of concomitant treatments. These inter-patient differences account for the difficulty in formulating subgroups using only several clinical parameters and biomarkers. In addition, even though the classification of a patients' (universal) clinical response to therapy can be standardized, the classification is subject to heterogeneity in response patterns over time (early response, late response, loss of response etc.), rapidly fluctuating disease activities and subjective influences. Also, the DAS28 has the disadvantage that it omits the assessment of regularly affected joints, such as those in the foot. The majority of the patients with a tender/swollen joint score of zero at baseline had involvement of the feet that prompted the prescription for a TNFi, which results in misclassification of these patients in categories with low disease activity and limits the ability to detect a EULAR response. Clinical and biological heterogeneity and (mis)classification of response makes the quest for biomarkers of response very challenging and will hamper future research on this topic, unless we find solutions for these issues.

## **CONCLUSION**

We showed that the combination of a proteomic score and clinical parameters is not able to identify TNFi non-responders and responders before the start of treatment. As earlier studies failed to produce reproducible results, it is questionable whether proteomics are useful in the prediction of response to TNFi treatment. Future studies on this topic should at least validate results in a separate cohort, to ensure the reproducibility and validity of the proposed proteomic predictor(s).

## REFERENCES

1. Wolfe AM, Kellgren JH, Masi AT. The epidemiology of rheumatoid arthritis: a review. II. Incidence and diagnostic criteria. *Bull Rheum Dis* 1968;19 3: 524-9.
2. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res* 2002;4 Suppl 3: S265-S72.
3. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med* 1999;340 4: 253-9.
4. Keystone EC, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum* 2004;50 5: 1400-11.
5. Cuppen BV, Welsing PM, Sprengers JJ, Bijlsma JW, Marijnissen AC, van Laar JM et al. Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability. *Rheumatology (Oxford)* 2016;55 5: 826-39.
6. Tracey D, Klareskog L, Sasso EH, Salfeld JG, Tak PP. Tumor necrosis factor antagonist mechanisms of action: a comprehensive review. *Pharmacol Ther* 2008;117 2: 244-79.
7. de Jager W, Prakken BJ, Bijlsma JW, Kuis W, Rijkers GT. Improved multiplex immunoassay performance in human plasma and synovial fluid following removal of interfering heterophilic antibodies. *J Immunol Methods* 2005;300 1-2: 124-35.
8. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38 1: 44-8.
9. van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;41 10: 1845-50.
10. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162 1: W1-73.
11. Hubbard AE, Ahern J, Fleischer NL, Van der Laan M, Lippman SA, Jewell N et al. To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. *Epidemiology* 2010;21 4: 467-74.
12. Fabre S, Dupuy AM, Dossat N, Guisset C, Cohen JD, Cristol JP et al. Protein biochip array technology for cytokine profiling predicts etanercept responsiveness in rheumatoid arthritis. *Clin Exp Immunol* 2008;153 2: 188-95.
13. Hueber W, Tomooka BH, Batliwalla F, Li W, Monach PA, Tibshirani RJ et al. Blood autoantibody and cytokine profiles predict response to anti-tumor necrosis factor therapy in rheumatoid arthritis. *Arthritis Res Ther* 2009;11 3: R76.
14. Lequerre T, Jouen F, Brazier M, Clayssens S, Klemmer N, Menard JF et al. Autoantibodies, metalloproteinases and bone markers in rheumatoid arthritis patients are unable to predict their responses to infliximab. *Rheumatology (Oxford)* 2007;46 3: 446-53.
15. Ortea I, Roschitzki B, Ovalles JG, Longo JL, de IT, I, Gonzalez I et al. Discovery of serum proteomic biomarkers for prediction of response to infliximab (a monoclonal anti-TNF antibody) treatment in rheumatoid arthritis: an exploratory analysis. *J Proteomics* 2012;77: 372-82.
16. Trocme C, Marotte H, Baillet A, Pallot-Prades B, Garin J, Grange L et al. Apolipoprotein A-I and platelet factor 4 are biomarkers for infliximab response in rheumatoid arthritis. *Ann Rheum Dis* 2009;68 8: 1328-33.
17. Uno K, Yoshizaki K, Iwahashi M, Yamana J, Yamana S, Tanigawa M et al. Pretreatment Prediction of Individual Rheumatoid Arthritis Patients' Response to Anti-Cytokine Therapy Using Serum Cytokine/Chemokine/Soluble Receptor Biomarkers. *PLoS One* 2015;10 7: e0132055.
18. Visvanathan S, Rahman MU, Keystone E, Genovese M, Klareskog L, Hsia E et al. Association of serum markers with improvement in clinical response measures after treatment with golimumab in patients with active rheumatoid arthritis despite receiving methotrexate: results from the GO-FORWARD study. *Arthritis Res Ther* 2010;12 6: R211.

19. Wright HL, Bucknall RC, Moots RJ, Edwards SW. Analysis of SF and plasma cytokines provides insights into the mechanisms of inflammatory arthritis and may predict response to therapy. *Rheumatology (Oxford)* 2012;51 3: 451-9.
20. Centola M, Cavet G, Shen Y, Ramanujan S, Knowlton N, Swan KA et al. Development of a multi-biomarker disease activity test for rheumatoid arthritis. *PLoS One* 2013;8 4: e60635.
21. Bakker MF, Cavet G, Jacobs JW, Bijlsma JW, Haney DJ, Shen Y et al. Performance of a multi-biomarker score measuring rheumatoid arthritis disease activity in the CAMERA tight control study. *Ann Rheum Dis* 2012;71 10: 1692-7.
22. Curtis JR, van der Helm-van Mil AH, Knevel R, Huizinga TW, Haney DJ, Shen Y et al. Validation of a novel multibiomarker test to assess rheumatoid arthritis disease activity. *Arthritis Care Res (Hoboken)* 2012;64 12: 1794-803.
23. Hirata S, Dirven L, Shen Y, Centola M, Cavet G, Lems WF et al. A multi-biomarker score measures rheumatoid arthritis disease activity in the BeSt study. *Rheumatology (Oxford)* 2013;52 7: 1202-7.
24. Hambardzumyan K, Bolce R, Saevarsdottir S, Cruickshank SE, Sasso EH, Chernoff D et al. Pretreatment multi-biomarker disease activity score and radiographic progression in early RA: results from the SWEFOT trial. *Ann Rheum Dis* 2015;74 6: 1102-9.

### **Supplementary text 1:** Detailed section of statistical analysis.

For protein values out of range we imputed values based on  $\frac{1}{2}$  of the minimum value (if true value below lower detection limit) or  $\frac{1}{2}$  of the highest measured value (if true value above upper detection limit). We accounted for missing data on selected parameters by performing multiple imputation, taking into account as many predictors as possible including DAS28 at baseline and three months[1]. Missing data for baseline clinical parameters did not exceed 15%. For patients with a missing DAS28 at three months (n=4 in development, n=18 in validation) and subsequent missing response, we used multiple imputation only for patients that were considered missing at random i.e. missing erythrocyte sedimentation rate (n=4 in validation), missed visit/lost to follow-up (n=2 in development, n=7 in validation) or discontinued treatment before three months due to practical reasons (n=1 in validation). For patients that missed a visit due to hospitalization regarding high RA disease activity (n=1 in validation) or discontinued treatment before three months of therapy due to inefficacy (n=2 in development, n=5 in validation), we performed a non-responder imputation considering the very clear unfavorable outcome. Per imputed clinical parameter, we investigated whether the imputations were successful based on the variability as seen in the iteration plots. Because all predictive abilities were tested on multiple imputed datasets, all results were pooled using Rubin's rules[2].

### **References**

1. Moons KG, Donders RA, Stijnen T, Harrell FE, Jr. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol* 2006;59 10: 1092-101.
2. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*: John Wiley & Sons, Inc, 1987.

**Supplementary Table 1:** Inflammatory proteins selected for analysis using xMAP technology.

<b>Protein</b>	<b>Alternative name</b>
CCL1	I-309
CCL2	MCP-1
CCL3	MIP1A
CCL4	MIP1B
CCL7	MCP-3
CCL11	EOTAXIN
CCL17	TARC
CCL19	MIP3B
CCL20	MIP3A
CCL22	MDC
CCL27	C-TACK
sCD14	
CXCL4	PF4
CXCL5	ENA-78
CXCL8	IL-8
CXCL9	MIG
CXCL10	IP-10
CXCL13	BLC
GM-CSF	
IFN $\alpha$	
IFN $\beta$	
IFN $\gamma$	
IL-1 $\beta$	
IL-1RA	
sIL-1RI	
sIL-1RII	
IL-2	
sIL-2R $\alpha$	sCD25
IL-4	
IL-5	
IL-6	
IL-7	
IL-10	
IL-12	
IL-13	
IL-15	
IL-16	
IL-17A	
IL-18	
IL-21	
IL-22	
IL-23	
IL-29	
sLAIR-1	CD305
sLIGHT	CD258
LT $\alpha$	(TNF $\beta$ )
MIF	CD74-L
NGF	bNGF
sPD-1	CD279
TNF $\alpha$	
sTNF-RI	TNFRSF1A
sTNF-RII	TNFRSF1B
TSLP	

Abbreviations (in alphabetical order): BLC = B lymphocyte chemoattractant; CCL = C-C chemokine ligand; CD = Cluster of Designation / Classification Determinant; CTACK = Cutaneous T cell-attracting chemokine; CXCL = C-X-C chemokine ligand; ENA-78 = Epithelial-derived neutrophil-activating peptide 78; EOTAXIN = Eosinophil chemotactic protein; GM-CSF = Granulocyte-macrophage colony-stimulating factor; IFN = Interferon; IL = Interleukin; IP10 = Interferon-gamma inducible protein 10; LAIR = Leukocyte-associated immunoglobulin-like receptor; LIGHT = Homologous to lymphotoxin exhibits inducible expression and competes with HSV glycoprotein D for binding to herpesvirus entry mediator, a receptor expressed on T lymphocytes; LT $\alpha$  = lymphotoxin alpha; MCP = Monocyte chemoattractant protein; MDC = Macrophage-derived chemokine; MIF = Macrophage migration inhibitory factor; MIG = Monokine induced by gamma interferon; MIP = Macrophage inflammatory protein; NGF/bNGF = Nerve growth factor/ beta nerve growth factor; PD-1 = programmed cell-death-1; PF4 = platelet factor 4; s = soluble; TARC = thymus and activation regulated chemokine; TNF = Tumor necrosis factor; TNF-R= Tumour necrosis factor receptor; TSLP = Thymic stromal lymphopoietin.

**Supplementary Table 2:** coefficients of selected proteins in the protein score, predicting DAS28 at three months. All proteins with relative normal distribution (39) together with CRP and DAS28 were first put in a partial least squares analyses (first PLS) on DAS28 at three months, after which proteins were selected based on a variable importance in projection (VIP)-score  $\geq 1$  and factor loading of one of the components so that no more than 15 were selected, which turned out to be  $>0.135$ . Then these selected parameters were put in PLS again (second PLS) to calculate the score with the best fit to the DAS28 at three months. Shown are the selected proteins with their VIP scores and factor loadings per component (first PLS), and the coefficients of the score (second PLS).

Protein	Value in score	VIP score	Factor loading C1	Factor loading C2	Factor loading C3	Coefficients protein score
<b>Intercept</b>						4.675
<b>CRP</b>	Ln(mg/L)	0.8	0.081	-0.018	0.001	-0.195
<b>DAS28</b>	Score <sup>#</sup>	2.2	0.091	0.020	0.155	0.334
<b>CCL3</b>	Ln(pg/ml)	1.1	0.109	0.001	-0.159	-0.109
<b>CCL17</b>	Ln(pg/ml)	1.5	-0.215	-0.006	0.001	-0.329
<b>CCL19</b>	Ln(pg/ml)	1.7	-0.029	-0.136	0.009	-0.155
<b>CCL22</b>	Ln(pg/ml)	1.5	0.114	-0.128	0.255	-0.035
<b>IL4</b>	Ln(pg/ml)	1.1	0.052	-0.235	0.175	-0.060
<b>IL6</b>	Ln(pg/ml)	1.0	-0.080	-0.184	0.282	-0.067
<b>IL7</b>	Ln(pg/ml)	1.1	0.170	-0.157	-0.066	-0.561
<b>IL15</b>	Ln(pg/ml)	1.5	0.077	-0.196	0.133	0.523
<b>sCD14</b>	Ln(pg/ml)	1.2	0.077	-0.033	-0.385	-0.463
<b>sCD74</b>	Ln(pg/ml)	1.5	0.150	-0.198	0.186	0.200
<b>sIL1-RI</b>	Ln(pg/ml)	1.2	0.036	-0.215	0.140	0.252
<b>sTNF-RII</b>	Ln(pg/ml)	1.2	0.097	-0.012	0.144	0.837

<sup>#</sup> $0.56 * \text{sqrt}(\text{TJC}) + 0.28 * \text{sqrt}(\text{swollen28}) + 0.70 * \text{ln}(\text{ESR}) + 0.014 * \text{GH}$

ESR = erythrocyte sedimentation rate; GH = visual analogue scale general health; TJC = tender joint count; SJC = swollen joint count

**Supplementary Table 3:** Serum concentrations of selected proteins in the development and validation cohort as measured by xMAP technology. The shown concentrations are all in pg/ml serum, yet only for the measurement of sCD14 the serum was diluted 1:1000. The values have not been log transformed and have not been adjusted. The patients in the development cohort were include between October 2012 and March 2014 (n=65) and xMAP analyses were performed in October 2014. Patients in the validation cohort were included from June 2009 and October 2012 (n=185) and xMAP analyses were performed in June 2016. The generally lower serum protein concentrations in the validation cohort could be explained by a longer storage time and subsequent decay of proteins, as the storage time for development and validation was 0-2 years and 4-7 years, respectively.

Protein	Development cohort (n=65)		Validation cohort (n=185)	
	Median	IQR	Median	IQR
<b>CCL3</b>	30.0	25.2 - 44.8	18.2	10.5 - 41.1
<b>CCL17</b>	235.9	151.9 - 379.4	263.3	184.6 - 412.6
<b>CCL19</b>	67.2	29.4 - 116.0	37.0	21.7 - 70.5
<b>CCL22</b>	482.5	332.2 - 618.8	337.2	232.0 - 457.9
<b>IL4</b>	2.2	0.6 - 4.7	0.1	0.1 - 0.1
<b>IL6</b>	22.2	9.0 - 56.2	1.9	0.0 - 15.6
<b>IL7</b>	32.1	23.5 - 43.1	21.6	5.9 - 38.5
<b>IL15</b>	11.4	9.1 - 14.9	4.6	2.6 - 7.8
<b>sCD14</b>	2.8E+6	2.2E+6 - 3.46E+6	2.6E+6	2.2E+6 - 3.1E+6
<b>sCD74</b>	919.9	572.8 - 1545.6	3869.9	1864.7 - 7403.5
<b>sIL1-RI</b>	10.5	8.0 - 13.3	0.9	0.1 - 2.0
<b>sTNF-RII</b>	1483.9	1297.5 - 1768.9	1859.9	1454.7 - 2218.0

IQR = inter quartile range; for abbreviations of proteins see Supplementary Table 1.





## Chapter 8

### **Exploring the inflammatory metabolomic profile to predict response to TNF- $\alpha$ inhibitors in rheumatoid arthritis**

Bart V.J. Cuppen<sup>#</sup>, Junzeng Fu<sup>#</sup>, Herman A. van Wietmarschen, Amy C. Harms, Slavik Koval, Anne C.A. Marijnissen, Judith J.W. Peeters, Johannes W.J. Bijlsma, Janneke Tekstra, Jacob M. van Laar, Thomas Hankemeier, Floris P.J.G. Lafeber, and Jan van der Greef, on behalf of all Society for Rheumatology Research Utrecht investigators

<sup>#</sup> These authors contributed equally

## ABSTRACT

### Objective

In clinical practice, approximately one-third of patients with rheumatoid arthritis (RA) respond insufficiently to TNF- $\alpha$  inhibitors (TNFi). The aim of the study was to explore the use of a metabolomics to identify predictors for the outcome of TNFi therapy, and study the metabolomic fingerprint in active RA irrespective of patients' response.

### Methods

In the metabolomic profiling, lipids, oxylipins, and amines were measured in serum samples of RA patients from the observational BiOCURA cohort, before start of biological treatment. Multivariable logistic regression models were established to identify predictors for good- and non-response in patients receiving TNFi (n=124). The added value of metabolites over prediction using clinical parameters only was determined by comparing the area under receiver operating characteristic curve (AUC-ROC), sensitivity, specificity, positive- and negative predictive value and by the net reclassification index (NRI). The models were further validated by 10-fold cross validation and tested on the complete TNFi treatment cohort including moderate responders. Additionally, metabolites were identified that cross-sectionally associated with the RA disease activity score based on a 28-joint count (DAS28), erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP).

### Results

Out of 139 metabolites, the best-performing predictors were *sn1*-LPC(18:3- $\omega$ 3/ $\omega$ 6), *sn1*-LPC(15:0), ethanolamine, and lysine. The model that combined the selected metabolites with clinical parameters showed a significant larger AUC-ROC than that of the model containing only clinical parameters ( $p = 0.01$ ). The combined model was able to discriminate good- and non-responders with good accuracy and to reclassify non-responders with an improvement of 30% (total NRI=0.23) and showed a prediction error of 0.27. For the complete TNFi cohort, the NRI was 0.22. In addition, 88 metabolites were associated with DAS28, ESR or CRP ( $p < 0.05$ ).

### Conclusion

Our study established an accurate prediction model for response to TNFi therapy, containing metabolites and clinical parameters. Associations between metabolites and disease activity may help elucidate additional pathologic mechanisms behind RA.

## INTRODUCTION

Rheumatoid arthritis (RA) is a chronic, disabling disease that mainly affects the synovial joints. The disease is a multifactorial autoimmune disorder with a prevalence of 0.5-1% in industrialized countries[1-3]. Disease-modifying anti-rheumatic drugs (DMARDs) are the cornerstone of anti-inflammatory therapy in RA and can be divided into two categories: conventional synthetic DMARDs (csDMARDs) and biological DMARDs (bDMARDs)[4]. The csDMARDs are relatively cheap chemical agents consisting of small active-substance molecules and are used for decades in the treatment of RA, whereas bDMARDs, predominantly antibodies such as TNF- $\alpha$  inhibitors (TNFis), are expensive agents that target specific inflammatory pathways and have revolutionized the treatment options for RA patients since the first trial in 1993[5]. Despite the success of TNFis, a substantial proportion of patients (approximately 30-40%) responds insufficiently to these biological agents[6, 7]. At the initiation of TNFi therapy it is as yet impossible to distinguish future responders from non-responders, therefore, the only used treatment approach is by trial and error. This approach is inefficient because the clinical response can only be assessed after at least three months of treatment. Within this timeframe, non-responders might develop joint damage or may experience toxic side effects. In addition, an inefficient treatment increases healthcare costs due to intensive monitoring, more complications, higher morbidity, and medication costs. The challenge is therefore to identify responders and non-responders before initiation of TNFi treatment so that the decision can be guided and the most optimal agent can be selected for each patient. Many approaches have been explored, mostly by evaluation of clinical parameters, proteins or mRNA biomarker profiles, but none were thus far successful in such a way that they can be implemented in clinical practice[8].

Metabolomics is a rapidly developing approach in biomarker research, involving the measurement of a large number of small-molecule metabolites in biological fluids, tissues and cells. One major advantage is that it offers a characteristic profile of each patient from minimal amounts of sample. With high-throughput techniques, such as nuclear magnetic resonance (NMR) spectroscopy and liquid chromatography coupled to mass spectrometry (LC-MS), metabolite profiles in disease or therapeutic response to treatment can be

measured [9]. In this sense, metabolomics provides a novel perspective on the search of new disease biomarkers and drug targets.

Metabolomic approaches in RA have already contributed to the understanding of RA and its subtypes, as well as the effect of drug treatment [10]. However, there is a limited number of studies employing metabolomic profiling to predict patients' response to biological therapies. To our knowledge, only two previous studies have used metabolomics to predict the clinical response to TNFi, both by usage of the <sup>1</sup>H-NMR-based technique. Kapoor et al.[11] screened the urine metabolome of 16 RA patients and found that histamine, glutamine, xanthurenic acid, and ethanolamine predicted TNFi response. In addition, there was a significant correlation between baseline urine metabolic profile and the magnitude of the one-year change in the disease activity. Priori et al.[12] showed that the serum metabolic profiling of 27 RA patients at baseline could discriminate the response to etanercept. Additionally, higher levels of isoleucine, leucine, valine, alanine, glutamine, tyrosine, and glucose and lower levels of 3-hydroxybutyrate were observed in good responders after 6 months of therapy. The predictors found in these <sup>1</sup>H-NMR based studies have thus far not been validated in other cohorts.

In the present study, the baseline serum metabolome of RA patients commencing biological therapy was analyzed using LC-MS, a different technique as previous studies used, targeting a large scope of metabolites consisting of lipids, oxylipins, and amines. The objective was to assess the potential value of serum metabolite profiles in the prediction of response to TNFi treatment, using LC-MS. Additionally, we investigated the association between metabolites with current disease activity, in order to gain more insight into the pathologic metabolic mechanisms in RA.

## **METHODS**

### **Patient cohort**

Patients were selected from the observational BiOCURA study (Biologicals and Outcome, Compared and predicted in Utrecht region, in Rheumatoid Arthritis) in which patients were enrolled between 2009 and 2015. In BiOCURA, RA patients eligible for bDMARD treatment in clinical practice, were followed up after start of treatment, in one academic

hospital and seven regional hospitals in the Netherlands (see Acknowledgements). The treatment included any of five TNFis, adalimumab, etanercept, infliximab, golimumab, and certolizumab pegol, or non-TNFi agents, including tocilizumab, abatacept, and rituximab. All csDMARDs were allowed to be used concomitantly with the bDMARD, and included methotrexate (MTX), hydroxychloroquine (HCQ), leflunomide (LEF), and glucocorticoids (GCs). Apart from csDMARDs, patients continued other medication, such as statins, bisphosphonates, anti-hypertension and nonsteroidal anti-inflammatory drugs (NSAIDs), all according to regular clinical practice.

Regular visits with the treating physician and a trained research nurse were scheduled at baseline, three, six, and twelve months. Clinical parameters and blood samples were collected by the nurse from each patient before the first dose of the biological agent. Of note, blood samples were collected from fasting and non-fasting patients, as the visiting times of the patients could not be standardized in the morning. The blood was collected in a Vacutainer® SST II tube and processed immediately after clotting. Samples were centrifuged for 10 min at 1500 *g* at room temperature and serum was aliquoted and stored at -80°C until use for metabolomic analyses. Re-inclusion after switching to a different biological agent was possible. The study was approved by the ethics committee of the UMC Utrecht and the institutional review boards of the participating centers (see Acknowledgements). Written informed consent was obtained from each patient.

Inclusion in the present study was restricted to subjects of BIOCURA fulfilling the following criteria: at start of treatment patients should not be in clinical remission (disease activity score based on a 28-joint count, DAS28 > 2.6), after three months of therapy the DAS28 assessment needed to be available, and no (temporary) discontinuation of treatment should have occurred within the first three months of bDMARD treatment.

### **Clinical measurements**

Demographic, clinical, and laboratory parameters of patients at baseline were obtained, including age, gender, menopausal status, body mass index (BMI), disease duration, any previously used bDMARD (biological naivety), currently used csDMARDs and non-anti-rheumatic drugs, 28 tender joint count (TJC), 28 swollen joint count (SJC), a 100mm visual analogue scale on general health (VAS-GH), erythrocyte sedimentation rate (ESR), C-

reactive protein (CRP), rheumatoid factor (RF), and anti-citrullinated protein antibody (ACPA). Disease activity was assessed at baseline and at follow-up visits, using DAS28 [13]. In clinical practice the response to biological therapy is usually measured 3-6 months after initiation [14]. However, in BiOCURA a substantial number of patients withdrew treatment before the 6-month time-point due to insufficient response or side effects. Using the 6-month response would thus result in (non-random) missing responses. Therefore, in this study, the patients' response was determined after 3-month of treatment, based on the EULAR response criteria[15]. A EULAR good response is defined as an improvement in DAS28 of  $> 1.2$  and a present DAS28  $\leq 3.2$ , whereas a EULAR non-response is assigned to patients with an improvement of 0.6-1.2 with present DAS28  $> 5.1$  or patients with an improvement  $\leq 0.6$ . In between, an improvement  $> 1.2$  with present DAS28  $> 3.2$  or an improvement of 0.6-1.2 with present DAS28  $< 5.1$  is specified as a EULAR moderate response.

### **Metabolomic profiling**

Serum samples from selected subjects were measured on three targeted LC-MS platforms, which used standard operating procedures from previously published methods [16-18], covering a broad spectrum of pre-defined metabolites. The lipid platform targets low abundance lipid species, including free fatty acids (FAs) and phospholipid derivatives, such as lysophosphatidylcholines (LPCs) and lysophosphatidylethanolamines (LPEs); the oxylipins platform covers oxygenated metabolites derived from polyunsaturated fatty acids through enzymatic and non-enzymatic oxidation processes; the amine platform targets amino acids and biogenic amines. All analyses were performed by the Biomedical Metabolomics Facility Leiden (BMFL) of the Leiden University. Extra serum of the subjects was pooled and used to create internal quality control (QC) samples.

### **Lipids analysis**

For the detection of lipids, each 20  $\mu\text{L}$  serum aliquot was spiked with internal standard (ISTD) mix and lipids were extracted by methanol. This lipid profiling was conducted using

ultra performance liquid chromatography coupled to electrospray ionization-quadrupole time-of-flight (Agilent 6530 San Jose, CA, USA) with an ACQUITY UPLC™ HSS T3 column (1.8 μm, 2.1×100mm) [16].

### **Oxylipins analysis**

Each 180 μL serum aliquot was spiked with antioxidants and ISTD mix, followed by solid phase extraction. The samples were analyzed by high-performance liquid chromatography (Agilent 1260, San Jose, CA, USA) coupled to a triple quadrupole mass spectrometer (Agilent 6460, San Jose, CA, USA), using an Ascentis® Express column (2.7 μm, 2.1x150 mm) [17].

### **Amines analysis**

Each 5 μL serum aliquot was spiked with an ISTD mix and proteins were precipitated by methanol, after which the supernatant was dried and derivatized by AQC reagent. The samples were analyzed by an ACQUITY ultra-performance liquid chromatography system coupled to Xevo Tandem quadrupole mass spectrometer (Waters, Milford, MA, USA) with an AccQ-Tag™ Ultra column (1.7 μm, 2.1x100 mm)[18].

### **Data preprocessing and correction**

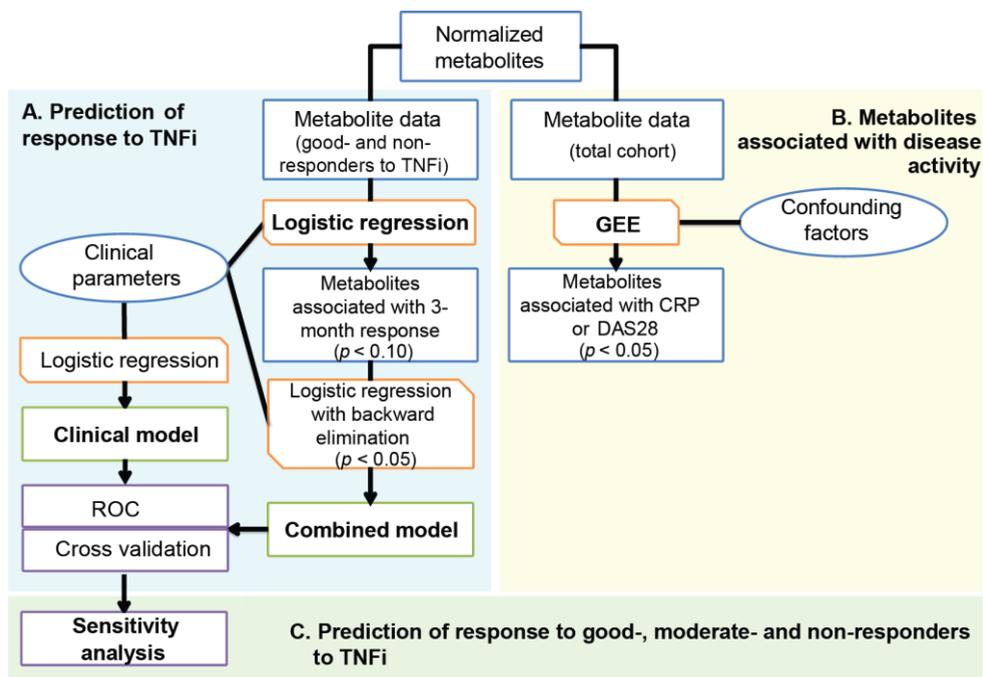
For the lipid and oxylipin platforms, peak determination and peak area integration were performed by Mass Hunter Quantitative Analysis (version B.05.00, Agilent technologies); for the amine platform, TargetLynx software (version 4.1, Waters) was employed. For each metabolite, the concentration was determined by the ratio between the peak area of targeted analyte and peak area of the appropriate ISTD. These response ratios ( $\text{Area}_{\text{analyte}}/\text{Area}_{\text{ISTD}}$ ) were used as raw metabolomic data in the subsequent analysis. For batches measured in each platform, an in-house developed method was applied to compensate and correct for instrumental drift during the measurements. The within-batch and between-batch effects of metabolomic data were corrected per metabolite using the responses of QC samples [19]. The QC samples were measured repeatedly every 10 patient samples across the different batches per platform. The details of the data correction method are described by van der Kloet *et al* [20]. The relative standard

deviation (RSD) of metabolites in the QC samples was used to assess the quality of targeted metabolites in each analytical platform.

Metabolite measurements, which were lower than the limitation of detection (LOD) of the platform, were imputed by half of the observed minimum value for the corresponding metabolite. Subsequently, log-transformation and auto-scaling of the metabolites were applied to reach or approximate a normal distribution, with a mean of 0 and SD of 1. The resulting data were used as input for all subsequent statistical analyses.

### Statistical analyses

Metabolites were used to develop a model for the prediction of response to TNFi at 3 months, and to assess their association with disease activity in general of the total cohort. An overview of the analyses is provided in Figure 1 and will be discussed in more detail below. All analyses were performed in IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp., Armonk, N.Y., USA), MedCalc for Windows, version 16.2.1 (MedCalc Software, Ostend, Belgium), and R (Version 3.2.3).



**Figure 1.** Flowchart of statistical analyses. (A) Prediction of response to TNFi: All steps to build a prediction model on TNFi response were performed on the TNFi subset with EULAR good-response or non-response (n=124). (B) Metabolites associated with disease activity. Analyses to investigate metabolites association with CRP, ESR or DAS28 were performed on the total cohort of patients using bDMARDs (n=231; including TNFi and non-TNFi treated patients). Blue boxes/circles indicate (selection of) respectively metabolites or clinical parameters, whereas orange boxes indicate the performed analyses. bDMARDs, biological disease-modifying anti-rheumatic drugs; CRP, C-reactive protein; DAS28, disease activity score based on a 28-joint count; ESR, erythrocyte sedimentation rate; GEE, generalized estimating equation, LC-MS, liquid chromatography coupled to mass spectrometry; ROC, receiver operating characteristic; TNFi, TNF- $\alpha$  inhibitor.

### **Development of models for predicting TNFi response**

In order to increase the possibility of picking up predictors related to response, we focused on subjects with EULAR good- *versus* non-response to TNFi treatment only, since EULAR moderate responders are an in-between category in which some patients continue treatment and others discontinue due to inefficacy (BiOCURA data). With TNFi good responders and non-responders as an outcome, we built two multivariable logistic regression models: a clinical model including clinical parameters only and one combined model with the same clinical parameters and the metabolites. We selected the following baseline clinical parameters: age, gender, menopausal status, BMI, smoking status, alcohol consumption, DAS28, (log-transformed) CRP, concomitant csDMARDs (MTX, HCQ, LEF, GCs) and non-DMARDs (statins, antihypertensive drugs, bisphosphonates and NSAIDs), regardless of the predictive ability (i.e. p-value) of each single parameter.

To decrease the number of metabolites in the combined model, a (non-strict) pre-selection on metabolites was performed. The predictive value of each individual metabolite to TNFi treatment outcome (good- or non-response) was investigated while combined with the clinical parameters, and all metabolites with  $p < 0.10$  were pre-selected. Subsequently, the pre-selected metabolites together with the clinical parameters were added to the logistic regression model on response with backward elimination of the metabolites, to build the final combined model (thus keeping all clinical parameters in).

The receiver operating characteristic (ROC) curves of the clinical and combined model was plotted, with the area under the curve (AUC) as an indicator of its predictive ability. Sensitivities, specificities, misclassification rates (MR), and positive- and negative predictive values (PPV/NPV) were calculated based on the optimal cutoff (Youden's index [21]). This optimal cut-off was applied to both models, so that per patient per model, either a non-response or good response was predicted, which could be compared to the observed response. Based on these predicted and observed responses, the net reclassification index (NRI) was calculated to determine if the addition of metabolites reclassifies more patients into the correct category. For example, when a future non-responder is classified at baseline to have a high probability of good response by the clinical model, but is classified into the low probability category in the combined model, the reclassification is "more correct" and the NRI for non-responders will increase, as well as the total NRI (=NRI responders + NRI non-responders). The robustness of the final models was judged by 10-fold cross validation of the model, using the cv.glm function of the R-package boot.

As an additional sensitivity analysis, the developed models were also applied to the complete cohort of TNFi initiating patients, thus without excluding any responders. In this step, the regression coefficients of the developed clinical and combined model were frozen and used to create a prediction rule. Subsequently, this clinical and combined prediction rule were compared for their abilities to distinguish EULAR non-responders from EULAR moderate- and good responders, using the same outcome measures for predictive ability as described before.

### **Metabolites and disease activity**

In order to investigate associations between metabolites and disease activity (CRP, ESR and DAS28), we analyzed the complete cohort of 231 patients with TNFi and non-TNFi therapy. However, in BiOCURA some RA patients were re-included after switching to a different biological agent (usually non-TNFi after a TNFi) and, therefore, had multiple baseline visits and follow-up periods. In order to account for the effects of subjects with multiple inclusions, generalized estimating equation (GEE) was used for these analyses, as GEE is a regression-based method that allows analyses of repeated measurements within

subjects [22]. Because we were interested in the association of each individual metabolite with CRP, ESR and DAS28, and clinical characteristics of patients might influence the metabolite levels (confounding), we corrected for baseline clinical parameters except CRP, ESR and DAS28, to gain more reproducible outcomes. As such, with the (log-transformed) CRP, (log-transformed) ESR or DAS28 as the dependent variable, each individual metabolite was added into GEE as an independent variable while corrected for the possible influential factors. Cytoscape was used to visualize the significant associations [23].

## **RESULTS**

In total, 231 RA patients from BiOCURA cohort fulfilled the selection criteria for the present study. Baseline characteristics of all patients are shown in Supplementary Table 1. Of all patients, 173 (74.9%) received TNFi treatment and 58 patients (25.1%) received a non-TNFi treatment. Concomitant csDMARDs and non-DMARDs were very diverse, and frequently included MTX (166 patients; 72%), GCs (93 patients; 40%) and bisphosphonates (101 patients; 44%) (Supplementary Table 2). Only 14 patients (6.1%) used no concomitant csDMARDs. The baseline characteristics of patients receiving TNFi and comparisons between non- and good responders are shown in Table 1. Among these patients, 64 were EULAR non-responders and 60 were good-responders. The number of RF positive patients and the SJC was significantly higher in the good responders, besides there were no significant differences in other baseline clinical variables between good responders and non-responders. Missing data was present for < 5% for each of the variables.

**Table 1:** Baseline characteristics of all selected TNFi initiating subjects (n=173), and split for all EULAR non-responders (n=64) and good responders (n=60). Descriptive statistics are expressed as number (%) for dichotomized variables, and mean  $\pm$  standard deviation (SD) and median and interquartile range (IQR) for respectively normally and non-normally distributed variables. The p-value is calculated for the difference between responders and non-responders.

	Subjects with TNFi (n=173)	Non-responders (n=64)	Good responders (n=60).	p-value
Female, n (%)	130 (75.1)	50 (78.1)	43 (71.1)	0.41
Menopausal status of females, n				0.84
Pre-menopause (%)	40 (30.8)	16 (32.0)	16 (37.2)	
Post-menopause(%)	82 (63.1)	28 (56.0)	26 (60.5)	
Unknown (%)	8 (6.1)	6 (12.0)	1 (2.3)	
Age, years, mean (SD)	54.6 (12.4)	53.7 (13.2)	53.9 (11.7)	0.95
Disease dur., years, median (IQR)	6.0 (2.0-12.0)	5.0 (2.0-11.5)	5.5 (2.0-11.0)	0.95
Smoking, currently, n (%)	42 (24.3)	17 (26.6)	15 (25.0)	0.84
Alcohol, >7 units week, n (%)	31 (17.9)	8 (12.5)	12 (20.0)	0.24
BMI, kg/m <sup>2</sup> , mean (SD)	26.8 (5.0)	27.3 (5.2)	26.5 (5.2)	0.42
Positive RF, n(%) <sup>a</sup>	114 (65.9)	39 (60.9)	47 (78.3)	0.04
Positive ACPA, n(%) <sup>a</sup>	123 (71.1)	41 (64.1)	46 (76.7)	0.13
CRP, mg/dL, median (IQR)	6.0 (3.0-13.0)	5.0 (2.8-10.0)	8.0 (3.8-15.0)	0.12
Baseline DAS28, mean (SD)	4.5 (1.1)	4.3 (1.2)	4.6 (0.9)	0.14
TJC, median (IQR)	7.0 (2.0-13.0)	6.5 (1.3-13.8)	6.5 (3.0-11.0)	0.73
SJC, median (IQR)	2.0 (0.0-4.0)	1.0 (0.0-3.0)	2.0 (1.0-4.0)	0.02
ESR, mm/h, median (IQR)	20.5 (11.0-34.8)	18.5 (5.3-34.0)	18.0 (10.3-33.0)	0.65
VAS-GH, mean (SD)	56.5 (23.1)	56.3 (23.4)	56.0 (23.2)	0.94

<sup>a</sup> Seropositivity for RF and ACPA was determined by the hospitals using different measurement methods and cut-offs, according to their own laboratory standards. Therefore we were not able to show an exact titer.

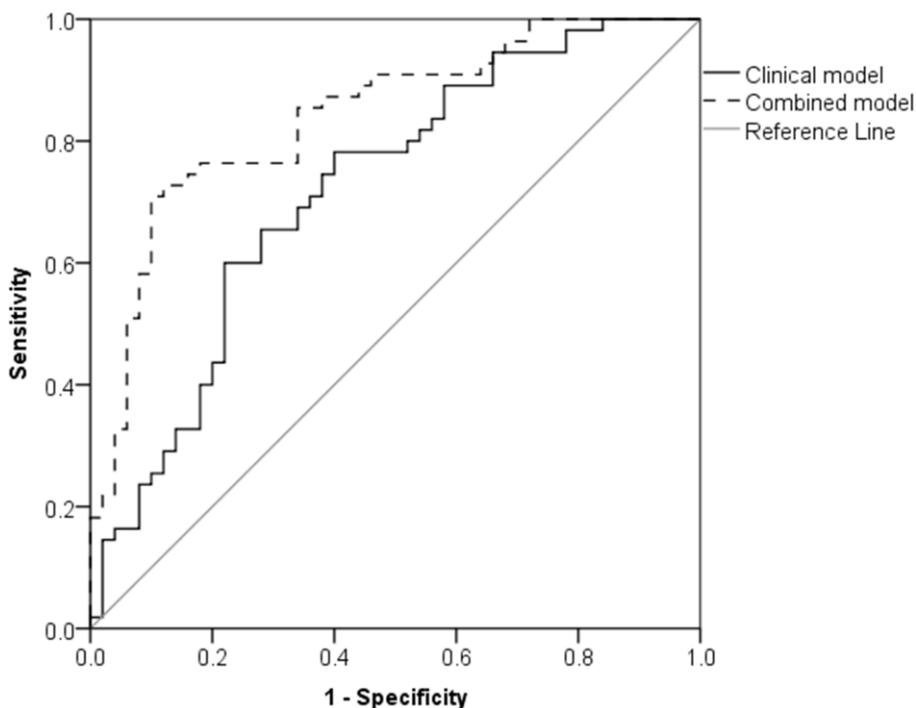
ACPA, anti-citrullinated protein antibody; BMI, body mass index; DAS28, disease activity score based on 28 joint count; ESR, erythrocyte sedimentation rate; IQR, interquartile range; RF, rheumatoid factor; SJC, 28 swollen joint count; TJC, 28 tender joint count; VAS-GH, 100mm visual analogue scale on general health.

The metabolites were measured using three validated platforms: lipids, oxylipins, and amines, thereby giving a broad view on the serum metabolites. 25 samples (of which 10 good-, 5 moderate-, and 10 non-responders) could not be analyzed in oxylipins platform due to inadequate serum volume. After analysis and data pre-processing, one sample in lipid platform and one sample in amine platform were excluded due to abnormal ISTD areas. For each platform, RSDs of QC samples were applied as quality indicators. In total, 139 metabolites were measured: 40 amines with QC RSD <0.15, 68 lipids (lysophospholipids and fatty acids) with QC RSD < 0.30, and 31 oxylipins both with QC RSD

< 0.30 (Supplementary Table 3e). The details of targeted compounds are shown in Supplementary Table 4-6.

### **Metabolites and TNFi response**

In the TNFi cohort, the clinical and combined model were built for the prediction of response containing all selected 16 clinical baseline parameters (as listed in the Materials and Methods section “Development of models for predicting TNFi response”). We did not exclude any clinical parameters, for three reasons. First, because the comparison of the clinical and combined model containing the same (complete) set of clinical parameters, allows us to compare the sole added value of metabolites over clinical parameters alone. Second, when biomarkers are in fact a (partial) reflection of any clinical parameter, we avoid the incorporation of any metabolite that provides knowledge which could have been obtained with a simple clinical parameter. Third, because clinical parameters might serve as partially influential factors (i.e. confounders) on metabolism as well, incorporation of all clinical parameter in the model would result in more reliable estimates of the true predictive ability of the metabolites and thus increases the reproducibility and externalization of results.



**Figure 2:** Receiver operating characteristic curves of clinical and combined model between good- and non-responders. The clinical model used 16 selected baseline clinical parameters; the combined model included four metabolites with  $p < 0.05$  in multivariable logistic regression with backward selection in addition to the clinical model.

The clinical model containing all selected 16 clinical baseline parameters (Cox & Snell R-square = 0.147, see Supplementary Figure 1) showed reasonable predictive abilities to differentiate good responders and non-responders, with an AUC-ROC of 0.720 (0.622-0.818) (Figure 2) and a sensitivity of 78.2%, specificity of 60.0%, PPV of 68.3%, NPV of 71.4% and MR of 30.5% (Supplementary Table 7).

Next, each metabolite was tested for their association with good and non-response to TNFi. *sn1*-LPC (18:3- $\omega$ 3/ $\omega$ 6), *sn1*-LPC (16:1), *sn1*-LPC (15:0), *sn2*-LPC (18:1), LPE (20:3- $\omega$ 3/ $\omega$ 6), LPE (18:1), *sn2*-LPC (18:0), LPE (18:0), *sn1*-LPC (18:0), 9,10-DiHOME, 9-HODE, 11-HDoHE, 8-HETE, 9,10-EpOME, ethanolamine, and lysine were preselected based on  $p <$

0.10. Multivariable logistic regression with backward selection was carried out on preselected metabolites to establish the final combined model. Four metabolites significantly added to the prediction of clinical response to TNFi therapy (Table 2).

**Table 2:** Remaining metabolites and their estimated contribution in the prediction of response to TNFi in the final prediction model.

	<b>Coefficient</b>	<b>Standard error</b>	<b>p-value</b>	<b>aOR (95%-CI)</b>
<b>sn1-LPC (18:3-ω3/ω6)</b>	-1.54	0.53	0.004	0.21 (0.08-0.61)
<b>sn1-LPC (15:0)</b>	1.67	0.56	0.003	5.32 (1.76-16.07)
<b>Ethanolamine</b>	-1.61	0.53	0.002	0.20 (0.07-0.57)
<b>Lysine</b>	1.02	0.40	0.010	2.78 (1.27-6.09)

aOR, adjusted odds ratio; CI, confidence interval.

**Table 3:** Net reclassification index of prediction models for good- and non-responders. An optimal cut-off for the clinical and combined model was chosen based on the Youden’s index, after which the predicted response of each patient per model was compared to the observed response. Shown are the number of patients, split for future non-responders and good responders (observed response), that were allocated at baseline to a predicted category (non-response/good-response) by both the clinical and combined model. These allocations could be correct or wrong, based on the observed response. There are four possibilities of allocations that represent an equally good or bad performance of both models (e.g. “30” represents 30 non-responders that were correctly classified as non-responders by both the clinical and combined model). Two categories denote an improvement in the prediction by the combined model: either a future non-responder that switches from response in the clinical model to non-response in the combined model (n=15), or a future responder switching from non-response to response (n=6). The two remaining discordant categories denote a worsening of prediction by the combined model. The NRI for non-responders was  $15/50 - 0/50 = 30\%$  improvement, while the NRI for responders was  $6/55 - 10/55 = -7\%$  due to a net worsening in prediction by the combined model. The total NRI was  $0.30 + (-0.07) = 0.23$ .

<b>Observed response (n=105)<sup>a</sup></b>	<b>Predicted by clinical model</b>	<b>Predicted by combined model</b>	
		<b>Non-response</b>	<b>Good response</b>
<b>Non-responders (n=50)</b>	<b>Non-response</b>	30 (equal)	0 (worsening)
	<b>Good response</b>	15 (improvement)	5 (equal)
<b>Good responders (n=55)</b>	<b>Non-response</b>	6 (equal)	6 (improvement)
	<b>Good response</b>	10 (worsening)	33 (equal)

<sup>a</sup> Due to the missing data of the clinical parameters, 19 out of 124 patients initiating TNFi therapy were excluded from the multivariable logistic regression models (clinical and combined model). Thus, 105 patients remained in the analyses.

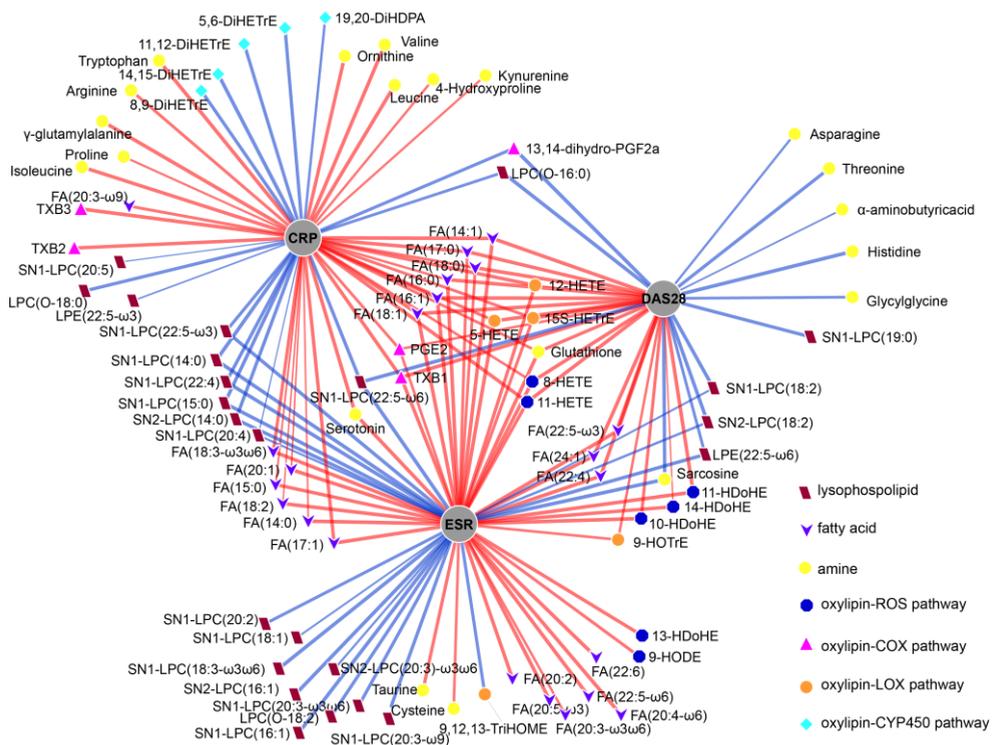
At baseline, good responders showed higher levels of *sn1*-LPC (15:0) and lysine, as well as lower levels of *sn1*-LPC (18:3- $\omega$ 3/ $\omega$ 6) and ethanolamine. The combined model (Cox & Snell R-square = 0.433, see Supplementary Figure 2) showed a good overall discriminative ability with an AUC-ROC of 0.841 (0.765-0.917) (Figure 2), and its performance is significantly better than the clinical model (difference between areas = 0.121,  $p = 0.01$ ). After determining the optimal cut-off, sensitivity was 70.9%, specificity 90.0%, PPV 88.6%, NPV 73.8% and MR 20.0% (Supplementary Table 7). The high specificity and PPV suggest that the patients classified as good responders are frequently correctly classified (low false positive rate).

Additionally an NRI was calculated in order to investigate whether the higher accuracy also results in a better reclassification of individual patients. The reclassification of non-responders by the combined model was 30% better and of responders 7% worse, resulting in a total NRI of 0.23 (Table 3). The clinical and combined models showed prediction errors of 0.303 and 0.269 respectively, which can be considered moderately robust. In the sensitivity analysis on all responders to TNFi, the clinical model performed weakly with an AUC-ROC of 0.641 (0.548-0.734), as compared to the combined model with 0.760 (0.682-0.837) (Supplementary Figure 3 & 4). The reclassification of non-responders by the combined model was 30% better and for responders 8% worse (total NRI = 0.22, Supplementary Table 8).

### **Metabolites and disease activity**

Metabolites were investigated for their association with disease activity. In total, 88 metabolites out of 139 were significantly associated with baseline DAS28, ESR or CRP ( $p < 0.05$ ) (Supplementary Table 9). In Figure 3, the associations between metabolites and the three clinical parameters are visualized. With respect to the amines, glutathione was positively associated with all the three parameters; sarcosine was negatively associated with both DAS28 and ESR, whereas serotonin was positively associated with CRP and ESR. Besides from the overlapping amines, five amines were negatively associated with baseline DAS28 (histidine, threonine, glycylglycine, asparagine, and  $\alpha$ -aminobutyric acid), while ten amines showed a positive association with CRP (valine, leucine, tryptophan,  $\gamma$ -glutamylalanin, isoleucine, ornithine, arginine, 4-hydroxyproline, kynurenine, and proline)

and two amines a positive association with ESR (taurine, cysteine). Fatty acids with aliphatic tails of 14 to 24 carbons were cross-sectionally found positively associated with CRP, ESR and/or DAS28, while lysophosphatidylethanolamines (LPEs) and lysophosphatidylcholines (LPCs) overall showed negative correlations with these parameters. Oxylipins were divided into different classes by the pathways— the auto-oxidation pathway with reactive oxygen species (ROS) and the enzymatic pathways with cyclooxygenase (COX), lipoxygenase (LOX) and cytochrome P450 (CYP450). Specifically oxylipins synthesized by ROS (10-HDoHE, 11-HDoHE, 11-HETE, 14-HDoHE, 8-HETE, 13-HDoHE, 9-HODE), COX (PGE2, TXB1, TXB2, TXB3) and LOX (12-HETE, 15S-HETrE, 5-HETE, 9-HOTrE) were cross-sectionally positively associated with CRP, ESR or DAS28, whereas only two downstream products in COX and LOX pathways, 13,14-dihydro-PGF2a and 9,12,13-TriHOME respectively, were found to be negatively associated with the clinical parameters. CYP450 synthesized oxylipins (11,12-DiHETrE, 14,15-DiHETrE, 19,20-DiHDPA, 5,6-DiHETrE, 8,9-DiHETrE) were all negatively associated with CRP.



**Figure 3:** Visualization of the associations between metabolites and disease activity – general inflammation (log-transformed CRP and ESR) and RA-specific inflammation (DAS28) – based on the complete cohort of bDMARD users (n=231). The metabolites that associated with either CRP, ESR or DAS based on linear generalized estimating equations (GEE), were grouped according to metabolic classes (LPCs, FAs, amines and oxylipins), which are represented as color-coded symbols adjacent to the metabolites. The metabolites in these metabolic classes showed comparable associations with CRP, ESR and/or DAS28. FAs positively- and the lysophospholipids negatively associated with CRP, ESR and/or DAS28; the association between other the oxylipins and amines with CRP, ESR and/or DAS28 were mixed, based on their metabolic functions. Positive associations are indicated with red lines, negative associations with blue lines; thicker lines indicate a more significant association. CRP, C-reactive protein; DAS28, disease activity score based on 28 joint counts; ESR: erythrocyte sedimentation rate; FA, fatty acid; LPE, lysophosphatidylethanolamine; LPC, lysophosphatidylcholine.

## DISCUSSION

One aim of the study was to explore the possibility of using baseline metabolomic profiling in the prediction of good- and non-response to TNFi treatment in patients with RA. Our combined model was able to predict response after 3 months of treatment with high

accuracy (AUC-ROC 0.841) and with a moderate robustness (prediction error: 0.269), and was significantly better than a model containing clinical parameters alone (increase in AUC-ROC = 0.121,  $p = 0.01$ , NRI = 0.23). The metabolites contributing to baseline prediction were *sn1*-LPC (15:0), *sn1*-LPC (18:3- $\omega$ 3/ $\omega$ 6), ethanolamine, and lysine, and the high NRI for non-responders indicates that these metabolites are especially useful to identify non-responders in advance. Additionally, our sensitivity analysis confirmed the added effect of metabolites in the prediction of (non-)response in all patients initiating TNFi. When the prediction rule is further validated, non-responders can be identified and offered more suitable treatments, which prevents joint damage, potentially toxic side-effects and saves healthcare costs.

Two of the identified predictors were also found to predict response in the two previous studies [11, 12]. A lower level of ethanolamine was also associated with response to TNFi in the study by Kapoor et al.[11], however, opposite results for a higher level of lysine were found, which predicted good response in our study and non-response in the study by Priori et al.[12]. In addition, both of these previous studies reported high values of glutamine as predictors of response to TNFi, though glutamine did not predict response in our study with- or without correction for multiple clinical parameters ( $p = 0.312$  and  $p = 0.555$  respectively). Several differences in study design compared to our study might explain these apparent discrepancies, such as different analytical platforms (technique and targeted panels) and statistical method applied, on top of a smaller sample size as well as the use of urine instead of serum in one of the two studies [12]. Therefore, *sn1*-LPC (15:0), *sn1*-LPC (18:3- $\omega$ 3/ $\omega$ 6), ethanolamine and lysine hold potential in predicting response to TNFi therapy when measured with LC-MS and on top of prediction by clinical parameters.

Two amino acids, ethanolamine and lysine, predicted good response to TNFi in our study. Ethanolamine is a primary amino acid and down-regulated levels were found to be predictive for good response to TNFi in our study. It has been reported that abundance of ethanolamine was lower in synovial fluid of RA patients than in non-RA patients [24]. The derivatives of ethanolamine play important roles in many pathways. For example, cytidine diphosphate -ethanolamine can be used as a substrate for *de novo* synthesizing phospholipids phosphatidylethanolamines (PE), which are structural components of

biological membranes [25]. Up-regulation of lysine was found to be predictive for good response to TNFi in the present study. Lysine is known to be related to RA treatment. A previous study that measured metabolites in 20 patients before and after three months of TNFi treatment, reported that the concentration of lysine was elevated after three months of TNFi therapy [26]. What its exact role in inflammation is remained speculative, however in animal studies, in which acute inflammatory processes were induced, it was indicated that amino acids are redirected from muscle to the liver for acute-phase protein synthesis and gluconeogenesis[27, 28]. Although absolute levels of ethanolamine and lysine were not directly found associated with DAS28 or CRP, it is possible that the relative abundance of these metabolites is to some extent informative for the inflammatory status in RA.

Interestingly, a down-regulation of *sn1*-LPC (18:3- $\omega$ 3/ $\omega$ 6) was found in good responders, while a down-regulated *sn1*-LPC (15:0) was found in non-responders. The most common structure of LPCs is with an even-chain fatty acid on *sn1* or *sn2* position, such as *sn1*-LPC (18:3- $\omega$ 3/ $\omega$ 6) with an 18-carbon fatty acid. However, *sn1*-LPC (15:0) carries an odd-saturated fatty acid (OCS-FA) chain. As part of lipid metabolism, LPCs and FAs can interconvert [29]. It is reported that the serum levels of OCS-FAs are associated with the consumption of dairy products and the reduced disease risks of coronary heart disease and type 2 diabetes [30], but no study has specified the function of OSC-FA chain LPCs in the human body, and the associations between LPCs and RA have not been explored. Although in the prediction models we adjusted the metabolites for multiple clinical parameters, we were unable to correct for the dietary influences or fasted/non-fasted status. We assumed that dietary variability would be equally distributed across response groups, however, the observed levels of OCS-FAs as predictors of response might indicate that this is not the case for dairy products.

Secondary aim of the present study was to explore the relation between metabolite profiles and RA disease activity in general while correcting for most important clinical parameters. We found 88 individual metabolites related to CRP, ESR or DAS28, among which LPCs and FAs were abundant. We found that LPCs were negatively associated with all three parameters, which is a consistent finding across several studies in humans[31, 32]. LPCs are lysophospholipids that play important roles in pathological processes as signaling molecules [33, 34]. The *sn1*-LPC is generated by the hydrolysis of

phosphatidylcholine (PC) by phospholipase A2 (PLA2). These LPCs have been studied extensively and their pro- or anti-inflammatory role and magnitude of effect are dependent on the length and (un)saturation of the fatty acyl group[35-37]. As the precursor of LPC, PC is present in the cell membrane and can bind to CRP to initiate host defense [38]. Long chain FAs are precursors of pro- and anti-inflammatory molecules[17, 39], which were found positively associated with CRP, ESR and DAS28 in this study. These results indicate that the FA metabolism was more active in the RA patients with higher disease activity and inflammation. Oxylipins generated by COX, LOX, and ROS pathway have potent pro-inflammatory effects and were positively associated with CRP or DAS28 while oxylipins derived from the CYP450 pathway were negatively associated with CRP, which can be explained by their anti-inflammatory effects [40].

In our study, several amines were found to be significantly associated with DAS28 (6 out of 7 negatively), CRP (all 12 positively) and ESR (4 out of 5 positively). Of those amines negatively related with DAS28, histidine, asparagine, and threonine were reported to be downregulated in RA patients compared to healthy controls [10, 26]. If these markers are indeed RA-specific, it is not surprising they also signify disease severity to a certain extent, which is indeed what can be concluded from the (negative) relation to DAS28 observed in our study. Amino acids in arginine metabolism (arginine and ornithine), tryptophan metabolism (serotonin and tryptophan) and branched-chain amino acids (isoleucine, leucine, and valine) were found to be positively associated with CRP, which may indicate as was suggested previously, that muscle proteins are degraded to amino acids and are redirected to the liver for acute-phase protein synthesis and gluconeogenesis[27, 28, 41, 42]. Serotonin and taurine, which were positively associated with CRP or ESR, were reported associated with oxidative stress and could therefore be linked to the inflammatory processes in RA [43, 44]. In conclusion, most of the metabolites we found to associate with disease inflammation have been described before, either in vivo or in vitro, and in most cases have been shown to be related to inflammation in RA in this study.

Although DAS28 has been extensively validated and is the most widely used instrument in clinical trials as well as in daily practice, it should also be noted that it has been claimed to be a rather instable monitoring instrument in RA patients with stable disease [45, 46]. Two recently developed and validated instruments –the simplified disease activity index (SDAI)

[14] and the clinical disease activity index (CDAI) [47], have been suggested to be used instead in clinical practice. Especially SDAI recently gained more relevance due to its inclusion in the American College of Rheumatism/EULAR remission criteria [48]. It has been reported that DAS28, SDAI, and CDAI do not result in the same classification of patients [49]. These newer instruments may reflect disease activity to a better extent, however SDAI and CDAI could not be used in our analyses because the Evaluator Global Assessment, a component included in both SDAI and CDAI, was not systemically collected in the BiOCURA study.

In this study, only metabolites at baseline were used to predict response and cross-sectionally investigate biomarkers for disease activity. As for the latter aim, investigating the change over time of metabolites might gain more insight in the most relevant metabolites regarding disease activity than any cross-sectional study, as they reflect prospective changes in RA patients.

## **CONCLUSION**

Metabolomic profiling is a powerful technique, which can be applied to analyze a wide range of metabolites from small sample volumes. Therefore, it has the potential of identification of biomarkers and increasing the understanding of the metabolic pathogenic pathways involved in a disease, such as RA. In the present study, we employed metabolic profiling to identify candidate metabolites to predict clinical response and assessed associations between metabolites and disease activity of RA. Because subjects were selected from an observational study, the heterogeneity needed to be adjusted for possible influential factors in all analyses. Yet, regardless of the possible absence of important influential factors, we showed that the predictive ability of a model can be quite high ( $AUC > 0.8$ ) in a heterogeneous setting like clinical practice. As for its potential use, metabolites allow a better identification of non-responders on top of clinical parameters, which is of added value in determining the most fitting treatment for each individual patient. Further external validation is needed to assess the robustness of these findings and its potential value for clinical application.

This is the first time that serum metabolomic profiles analyzed by LC-MS have been demonstrated to predict therapeutic response to biological treatment in RA. It would be worth studying how these metabolites could be used for predicting patients' response with external validation in a more homogenous cohort and thereby potentially optimize the treatment strategy for patients with RA.

## REFERENCES

1. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. *Arthritis research* 2002;4: S265-S72.
2. Wolfe AM, Kellgren JH, Masi AT. The epidemiology of rheumatoid arthritis: a review. II. Incidence and diagnostic criteria. *Bulletin on the rheumatic diseases* 1968;19: 524-9.
3. Gabriel SE, Michaud K. Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. *Arthritis research & therapy* 2009;11: 229.
4. Smolen JS, Landewé R, Breedveld FC, Dougados M, Emery P, Gaujoux-Viala C et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. *Annals of the rheumatic diseases* 2014;73: 492-509.
5. Elliott MJ, Maini RN, Feldmann M, Long-Fox A, Charles P, Katsikis P et al. Treatment of rheumatoid arthritis with chimeric monoclonal antibodies to tumor necrosis factor alpha. *Arthritis and rheumatism* 1993;36: 1681-90.
6. Keystone EC, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS et al. Radiographic, Clinical, and Functional Outcomes of Treatment with Adalimumab (a Human Anti-Tumor Necrosis Factor Monoclonal Antibody) in Patients with Active Rheumatoid Arthritis Receiving Concomitant Methotrexate Therapy: A Randomized, Placebo-Controlled. *Arthritis and Rheumatism* 2004;50: 1400-11.
7. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *The New England journal of medicine* 1999;340: 253-9.
8. Cuppen BVJ, Welsing PMJ, Sprengers JJ, Bijlsma JWJ, Marijnissen ACA, van Laar JM et al. Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability. *Rheumatology* 2016;55: 826-39.
9. Guma M, Tiziani S, Firestein GS. Metabolomics in rheumatic diseases: desperately seeking biomarkers. *Nature reviews Rheumatology* 2016;12: 269-81.
10. Priori R, Scrivero R, Brandt J, Valerio M, Casadei L, Valesini G et al. Metabolomics in rheumatic diseases: The potential of an emerging methodology for improved patient diagnosis, prognosis, and treatment efficacy. *Autoimmunity Reviews* 2013;12: 1022-30.
11. Kapoor SR, Filer A, Fitzpatrick Ma, Fisher Ba, Taylor PC, Buckley CD et al. Metabolic profiling predicts response to anti-tumor necrosis factor  $\alpha$  therapy in patients with rheumatoid arthritis. *Arthritis and Rheumatism* 2013;65: 1448-56.
12. Priori R, Casadei L, Valerio M, Scrivero R, Valesini G, Manetti C. (1)H-NMR-Based Metabolomic Study for Identifying Serum Profiles Associated with the Response to Etanercept in Patients with Rheumatoid Arthritis. *PLoS One* 2015;10 11: e0138537.
13. Prevoo M, van 't Hof M, Kuper H, van Leeuwen M, van de Putte L, van Riel P. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis & Rheumatism* 1995;38: 44-8.
14. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology* 2003;42: 244-57.
15. van Gestel A, Haagsma C, van Riel P. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis & Rheumatism* 1998;41: 1845-50.
16. Hu C, van Dommelen J, van der Heijden R, Spijksma G, Reijmers TH, Wang M et al. RPLC-Ion-Trap-FTMS Method for Lipid Profiling of Plasma: Method Validation and Application to p53 Mutant Mouse Model. *Journal of Proteome Research* 2008;7: 4982-91.
17. Strassburg K, Huijbrechts AML, Kortekaas Ka, Lindeman JH, Pedersen TL, Dane A et al. Quantitative profiling of oxylipins through comprehensive LC-MS/MS analysis: application in cardiac surgery. *Analytical and bioanalytical chemistry* 2012;404: 1413-26.
18. Noga MJ, Dane A, Shi S, Attali A, van Aken H, Suidgeest E et al. Metabolomics of cerebrospinal fluid reveals changes in the central nervous system metabolism in a rat model of multiple sclerosis. *Metabolomics* 2012;8: 253-63.
19. Draisma HH, Reijmers TH, van der Kloet F, Bobeldijk-Pastorova I, Spies-Faber E, Vogels JT et al. Equating, or correction for between-block effects with application to body fluid

- LC-MS and NMR metabolomics data sets. *Anal Chem* 2010;82 3: 1039-46.
20. van der Kloet FM, Bobeldijk I, Verheij ER, Jellema RH. Analytical error reduction using single point calibration for accurate and precise metabolomic phenotyping. *Journal of Proteome Research* 2009;8: 5132-41.
  21. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3: 32-5.
  22. Hubbard AE, Ahern J, Fleischer NL, Van der Laan M, Lippman Sa, Jewell N et al. To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. *Epidemiology* 2010;21: 467-74.
  23. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 2003;13: 2498-504.
  24. Kim S, Hwang J, Xuan J, Jung YH, Cha HS, Kim KH. Global metabolite profiling of synovial fluid for the specific diagnosis of rheumatoid arthritis from other inflammatory arthritis. *PLoS One* 2014;9 6: e97501.
  25. Gibellini F, Smith TK. The Kennedy pathway-De novo synthesis of phosphatidylethanolamine and phosphatidylcholine. *IUBMB Life* 2010;62 6: 414-28.
  26. Zabek A, Swierkot J, Malak A, Zawadzka I, Deja S, Bogunia-Kubik K et al. Application of (1)H NMR-based serum metabolomic studies for monitoring female patients with rheumatoid arthritis. *Journal of pharmaceutical and biomedical analysis* 2016;117: 544-50.
  27. Fischer JE, Hasselgren PO. Cytokines and glucocorticoids in the regulation of the "hepato-skeletal muscle axis" in sepsis. *The American Journal of Surgery* 1991;161: 266-71.
  28. Austgen TR, Chakrabarti R, Chen MK, Souba WW. Adaptive regulation in skeletal muscle glutamine metabolism in endotoxin-treated rats. *The Journal of trauma* 1992;32: 600-7.
  29. Shindou H, Shimizu T. Acyl-CoA:lysophospholipid acyltransferases. *The Journal of biological chemistry* 2009;284: 1-5.
  30. Jenkins B, West JA, Koulman A. A review of odd-chain fatty acid metabolism and the role of pentadecanoic acid (C15:0) and heptadecanoic acid (C17:0) in health and disease. *Molecules* 2015;20: 2425-44.
  31. Wallace M, Morris C, O'Grada CM, Ryan M, Dillon ET, Coleman E et al. Relationship between the lipidome, inflammatory markers and insulin resistance. *Molecular bioSystems* 2014;10: 1586-95.
  32. Heimerl S, Fischer M, Baessler A, Liebisch G, Sigrüener A, Wallner S et al. Alterations of Plasma Lysophosphatidylcholine Species in Obesity and Weight Loss. *PLoS ONE* 2014;9: e111348.
  33. Ye X. Lysophospholipid signaling in the function and pathology of the reproductive system. *Human Reproduction Update* 2008;14: 519-36.
  34. Ginsburg I, Ward PA, Varani J. Lysophosphatides enhance superoxide responses of stimulated human neutrophils. *Inflammation* 1989;13: 163-74.
  35. Huang YH, Schäfer-Elinder L, Wu R, Claesson HE, Frostegård J. Lysophosphatidylcholine (LPC) induces proinflammatory cytokines by a platelet-activating factor (PAF) receptor-dependent mechanism. *Clinical and Experimental Immunology* 1999;116: 326-31.
  36. Kabarowski JH. G2A and LPC: Regulatory functions in immunity. *Prostaglandins and Other Lipid Mediators* 2009;89: 73-81.
  37. Hung ND, Sok DE, Kim MR. Prevention of 1-palmitoyl lysophosphatidylcholine-induced inflammation by polyunsaturated acyl lysophosphatidylcholine. *Inflammation Research* 2012;61: 473-83.
  38. Volanakis J. Human C-reactive protein: expression, structure, and function. *Molecular Immunology* 2001;38: 189-97.
  39. Calder PC. Polyunsaturated fatty acids and inflammation. *Prostaglandins Leukot Essent Fatty Acids* 2006;75 3: 197-202.
  40. Heemskerck MM, Dharuri HK, van den Berg SA, Jonasdottir HS, Kloos DP, Giera M et al. Prolonged niacin treatment leads to increased adipose tissue PUFA synthesis and anti-inflammatory lipid and oxylipin plasma profile. *J Lipid Res* 2014;55 12: 2532-40.
  41. Nicastro H, da Luz CR, Chaves DF, Bechara LR, Voltarelli VA, Rogero MM et al. Does Branched-Chain Amino Acids Supplementation Modulate Skeletal Muscle Remodeling through Inflammation Modulation? Possible Mechanisms of Action. *J Nutr Metab* 2012;2012: 136937.
  42. Ouyang X, Dai Y, Wen J, Wang L. 1H NMR-based metabolomic study of metabolic profiling for systemic lupus erythematosus. *Lupus* 2011;20: 1411-20.
  43. Sugiura T, Dohi Y, Yamashita S, Hirowatari Y, Fujii S, Ohte N. Serotonin in peripheral

- blood reflects oxidative stress and plays a crucial role in atherosclerosis: Novel insights toward holistic anti-atherothrombotic strategy. *Atherosclerosis* 2016;246: 157-60.
44. Marcinkiewicz J, Kontny E. Taurine and inflammatory diseases. *Amino Acids* 2014;46 1: 7-20.
  45. Behrens F, Tony H-P, Alten R, Kleinert S, Scharbatke EC, Köhm M et al. Development and validation of a new disease activity score in 28 joints-based treatment response criterion for rheumatoid arthritis. *Arthritis care & research* 2013;65: 1608-16.
  46. Madsen OR. Stability of clinical outcome measures in rheumatoid arthritis patients with stable disease defined on the basis of the EULAR response criteria. *Clinical Rheumatology* 2016.
  47. Aletaha D, Nell V, Stamm T, Uffmann M, Pflugbeil S, Machold K et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Research and Therapy* 2005;7: R796-R806.
  48. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LHD, Funovits J et al. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis & Rheumatism* 2011;63: 573-86.
  49. Martins FM, da Silva JAP, Santos MJ, Vieira-Sousa E, Duarte C, Santos H et al. DAS28, CDAI and SDAI cut-offs do not translate the same information: results from the Rheumatic Diseases Portuguese Register Reuma.pt. *Rheumatology* 2015;54: 286-91.

**Supplementary Table 1:** Baseline characteristics of all selected subjects (n=231), and split for all EULAR good-responders and non-responders (n=80 each). Descriptive statistics are expressed as number (%) for dichotomized variables, and mean  $\pm$  standard deviation (SD) and median and interquartile range (IQR) for respectively normally and non-normally distributed variables. The p-value is calculated for the difference between responders and non-responders.

	All (n=231)	Non-resp. (n=80)	Good resp. (n=80)	p-value
Gender, female, n (%)	172 (74.5)	64 (80.0)	57 (71.3)	0.28
Menopausal status of females, n(%)				0.82
Pre-menopause	55 (23.8)	22 (34.4)	22 (38.6)	
Post-menopause	107 (46.3)	36 (56.3)	33 (57.9)	
Unknown	10 (4.3)	6 (9.4)	2 (3.5)	
Age, years, mean (SD)	54.5 (13.1)	52.8 (13.5)	52.8 (13.0)	1.00
Disease dur. years, mean (SD)	7.0 (3.0-14.0)	6.0 (3.0-12.0)	7.0 (3.0-12.8)	0.79
Smoking, currently	53 (22.9)	18 (22.5)	18 (22.5)	1.00
Alcohol, >7 units/week	39 (17.0)	9 (11.3)	16 (20.3)	0.13
BMI, kg/m <sup>2</sup> , mean (SD)	26.6 (5.1)	27.1 (5.1)	26.5 (4.9)	0.47
Positive RF, n(%)	151 (65.4)	47 (58.8)	59 (73.8)	0.07
Positive ACPA, n(%)	161 (69.7)	50 (62.5)	59 (73.8)	0.17
CRP, mg/dL, median (IQR)	6.0 (3.0-16.5)	5.0 (2.0-10.5)	8.0 (3.0-20.3)	0.05
Baseline DAS28, mean (SD)	4.7 (1.1)	4.4 (1.2)	4.6 (0.9)	0.18
TJC, median (IQR)	7.0 (2.0-14.0)	7.0 (2.0-14.0)	7.0 (3.0-12.8)	0.71
SJC, median (IQR)	2.0 (0.0-4.0)	1.0 (0.0-3.0)	2.0 (0.0-4.0)	0.06
ESR, mm/h, median IQR	21.0 (21.0-39.0)	18.0 (6.0-34.0)	18.0 (10.0-36.5)	0.53
VAS-GH, mean (SD)	58.2 (22.9)	58.2 (23.1)	55.8 (22.6)	0.50

ACPA, anti-citrullinated protein antibody; BMI, body mass index; DAS28, disease activity score based on 28 joint count; ESR, erythrocyte sedimentation rate; IQR, interquartile range; RF, rheumatoid factor; SJC, 28 swollen joint count; TJC, 28 tender joint count; VAS-GH, 100mm visual analogue scale on general health.

**Supplementary Table 2:** Previously and currently used treatments of all selected subjects (n = 231) and split for good responders and non-responders (n = 80 each). P-values were calculated based on chi-square/fisher exact.

Treatment	All (n=231)	Non-resp. (n=80)	Good resp. (n=80)	p-value
<b>No. of previously used bDMARDs, n (%)</b>				0.16
<b>0 (naïve)</b>	116 (50.2)	37 (46.3)	49 (61.3)	
<b>1</b>	56 (24.2)	24 (30.0)	17 (21.3)	
<b>&gt;1</b>	59 (25.5)	17 (23.8)	14 (17.5)	
<b>Initiated treatment, n (%)</b>				0.10
<b>Adalimumab</b>	70 (30.3)	25 (31.3)	27 (33.8)	
<b>Etanercept</b>	60 (26.0)	22 (27.5)	23 (28.8)	
<b>Golimumab</b>	23 (10.0)	7 (8.8)	5 (6.3)	
<b>Infliximab</b>	10 (4.3)	6 (7.5)	3 (3.8)	
<b>Certolizumab</b>	10 (4.3)	4 (5.0)	2 (2.5)	
<b>Abatacept</b>	19 (8.2)	8 (10.0)	3 (3.8)	
<b>Tocilizumab</b>	25 (10.8)	3 (3.8)	14 (17.5)	
<b>Rituximab</b>	14 (6.1)	5 (6.3)	3 (3.8)	
<b>Concomitant DMARDs, n (%)</b>	217 (93.9)	76 (95.0)	75 (93.8)	1.00
<b>Methotrexate</b>	166 (71.9)	55 (68.8)	64 (80)	0.15
<b>Hydroxychloroquine</b>	50 (21.6)	18 (22.5)	19 (23.8)	1.00
<b>Leflunomide</b>	28 (12.1)	11 (13.8)	5 (6.3)	0.19
<b>Glucocorticoids</b>	93 (40.3)	36 (45.0)	26 (32.5)	0.14
<b>Concomitant non-DMARDs, n (%)</b>				
<b>Statins</b>	34 (14.7)	14 (17.5)	6 (7.5)	0.09
<b>Anti-osteoporosis</b>	101 (43.7)	37 (46.3)	32 (40.0)	0.52
<b>Anti-hypertention</b>	68 (29.4)	25 (31.2)	17 (21.3)	0.21
<b>NSAIDs</b>	142 (61.5)	45 (56.3)	49 (61.3)	0.63

P-values were calculated based on chi-square/fisher exact.

DMARDs, biological DMARDs, DMARDs, disease-modifying antirheumatic drugs.

**Supplementary Table 3:** List of relative standard deviations (RSD) for all 139 measured metabolites. The RSDs were calculated on the metabolites in the QC samples. The "sn1-" or "sn2-" prefix is used to indicate the position of the fatty acid chain esterified to the glycerol backbone of a lysophospholipid. The 'O-' prefix is used to identify a plasmalogen lysophospholipid, where the fatty acid chain is attached via a vinyl ether linkage to the glycerol backbone. The 'ω3', 'ω6', 'ω9' are used to indicate double bond (C=C) at the third, sixth or ninth carbon atom from the end of the fatty acid chain.

Phospholipids*	RSD	Fatty acids*	RSD	Oxylipins	RSD	Amines	RSD
LPE (16:0)	5%	FA (14:0)	8%	9,10-DiHOME	6%	Ethanolamine	6%
LPE (18:0)	3%	FA (14:1)	26%	12,13-DiHOME	6%	Ornithine	3%
LPE (18:1)	5%	FA (15:0)	23%	TXB2	7%	Lysine	5%
LPE (18:2)	5%	FA (16:0)	7%	13-HODE	7%	Glycine	4%
LPE (20:3-ω3ω6)	8%	FA (16:1)	5%	14-HDoHE	8%	Sarcosine	10%
LPE (20:4)	5%	FA (17:0)	12%	13-HDoHE	8%	Alanine	3%
LPE (20:5)	10%	FA (17:1)	11%	14,15-DiHETrE	10%	3-Aminoisobutyric acid	6%
LPE (22:5-ω3)	9%	FA(18:0)	9%	13,14-dihydro-PGF2a	11%	α-aminobutyric acid	3%
LPE (22:5-ω6)	22%	FA (18:1)	3%	9-HODE	11%	Serine	4%
LPE (22:6)	4%	FA (18:2)	3%	19,20-DiHDPA	12%	Cystathionine	7%
sn2-LPC (14:0) <sup>a</sup>	11%	FA (18:3-ω3ω6)	3%	11-HETE	13%	Proline	14%
sn1-LPC (14:0) <sup>b</sup>	6%	FA (20:0)	18%	15S-HETrE	14%	Valine	4%
sn1-LPC (15:0)	7%	FA (20:1)	8%	10-HDoHE	15%	Threonine	4%
sn2-LPC (16:0)	7%	FA (20:2)	13%	5,6-DiHETrE	16%	Cysteine	6%
sn1-LPC (16:0)	2%	FA (20:3-ω3ω6)	6%	11,12-DiHETrE	16%	Taurine	4%
sn2-LPC (16:1)	8%	FA (20:3-ω9)	20%	LTB4	17%	Pipecolic acid	6%
sn1-LPC (16:1)	6%	FA (20:4-ω6)	4%	8,9-DiHETrE	17%	Isoleucine	3%
sn2-LPC (18:0)	3%	FA (20:5-ω3)	6%	9,10-EpOME	17%	Leucine	4%
sn1-LPC (18:0)	2%	FA (22:4)	6%	5-HETE	18%	4-Hydroxyproline	3%
sn2-LPC (18:1)	4%	FA (22:5-ω3)	3%	PGE2	18%	Asparagine	3%
sn1-LPC (18:1)	6%	FA (22:5-ω6)	4%	17,18-DiHETE	18%	Glycylglycine	4%
sn2-LPC (18:2)	8%	FA (22:6)	6%	12-HETE	18%	Aspartic acid	4%
sn1-LPC (18:2)	7%	FA (24:0)	15%	20-carboxy-LTB4	18%	s-Methylcysteine	5%

<i>sn2</i> -LPC (18:3- $\omega$ 3 $\omega$ 6)	20%	FA (24:1)	9%	TXB1	19%	Homocysteine	7%
<i>sn1</i> -LPC (18:3- $\omega$ 3 $\omega$ 6)	7%			12,13-EpOME	19%	O-Phosphoethanolamine	14%
<i>sn1</i> -LPC (19:0)	24%			9-HOTrE	20%	Glutamine	4%
<i>sn1</i> -LPC (20:1)	13%			11-HDoHE	20%	Glutamic acid	4%
<i>sn1</i> -LPC (20:2)	7%			9,12,13-TriHOME	20%	Methionine	4%
<i>sn2</i> -LPC (20:3- $\omega$ 3 $\omega$ 6)	9%			TXB3	21%	Histidine	5%
<i>sn1</i> -LPC (20:3- $\omega$ 3 $\omega$ 6)	9%			8-HETE	22%	Methionine sulfoxide	9%
<i>sn1</i> -LPC (20:3- $\omega$ 9)	17%			12S-HHTrE	30%	Phenylalanine	3%
<i>sn2</i> -LPC (20:4)	6%					1-Methylhistidine	15%
<i>sn1</i> -LPC (20:4)	8%					Arginine	8%
<i>sn2</i> -LPC (20:5)	9%					Citrulline	3%
<i>sn1</i> -LPC (20:5)	8%					Serotonin	11%
<i>sn1</i> -LPC (22:4)	9%					Tyrosine	4%
<i>sn1</i> -LPC (22:5- $\omega$ 3) <sup>d</sup>	8%					Tryptophan	4%
<i>sn1</i> -LPC (22:5- $\omega$ 6)	9%					Kynurenine	4%
<i>sn2</i> -LPC (22:6)	8%					$\gamma$ -glutamylalanine	12%
<i>sn1</i> -LPC (22:6)	7%					Glutathione	10%
LPC (O-16:0) <sup>c</sup>	6%						
LPC (O-18:0)	12%						
LPC (O-18:1)	6%						
LPC (O-18:2)	19%						

\*The number in brackets represent the long-chain fatty acid attached in the lysophospholipids (e.g. FA(14:0) represents a fatty acid with 14 carbons and 0 double bonds).

LPE, lysophosphatidylethanolamine; LPC, lysophospholipids; FA, fatty acid; LT<sub>B4</sub>, leukotriene B<sub>4</sub>; DiHDPA, dihydroxy-docosapentaenoic acid; DiHETE, dihydroxy-eicosatetraenoic acid; DiHETrE, dihydroxy-eicosatrienoic acid; DiHOME, dihydroxy-octadecenoic acid; PGF<sub>2</sub> $\alpha$ , prostaglandin F<sub>2</sub> $\alpha$ ; EpOME, epoxy-octadecenoic acid; HDoHE, hydroxyl-docosahexaenoic acid; HEPE, hydroxy-eicosapentaenoic acid; HETE, hydroxyicosatetraenoic acid; HETrE, hydroxy-eicosatrienoic acid; HHTrE, hydroxy-heptadecatrienoic acid; HODE, hydroxy-octadecadienoic acid; HOTrE, hydroxy-octadecatrienoic acid; KODE, oxo-octadecadienoic acid; TriHOME, trihydroxy-octadecenoic acid.

**Supplementary Table 4:** List of detected metabolites in lipids analysis

Lipid Class	Lipid Maps	Metabolite species	Amount (n)
Free fatty acids (FA)	FA01	14:0, 14:1, 15:0, 16:0, 16:1, 17:0, 17:1, 18:0, 18:1, 18:2, 18:3- $\omega$ 3 $\omega$ 6, 20:0, 20:1, 20:2, 20:3- $\omega$ 3 $\omega$ 6, 20:3- $\omega$ 9, 20:4- $\omega$ 6, 20:5- $\omega$ 3, 22:4, 22:5- $\omega$ 3, 22:5- $\omega$ 6, 22:6, 24:0, 24:1	24
Lysophosphatidylcholine (LPC)	GP0105	<i>sn1</i> : , 14:0, 15:0, 16:0, 16:1, 18:0, 18:1, 18:2, 18:3- $\omega$ 3 $\omega$ 6, 19:0, 20:1, 20:2, 20:3- $\omega$ 3 $\omega$ 6, 20:3- $\omega$ 9, 20:4, 20:5, 22:4, 22:5- $\omega$ 3, 22:5- $\omega$ 6, 22:6 <i>sn2</i> : , 14:0, 16:0, 16:1, 18:0, 18:1, 18:2, 18:3- $\omega$ 3 $\omega$ 6, 20:3- $\omega$ 3 $\omega$ 6, 20:4, 20:5, 22:6	30
Lysophosphatidylethanolamine (LPE)	GP0205	16:0, 18:0, 18:1, 18:2, 20:3- $\omega$ 3 $\omega$ 6, 20:4, 20:5, 22:5- $\omega$ 3, 22:5- $\omega$ 6, 22:6	10
Plasmalogen Lysophosphatidylcholine (pLPC)	GP0106	C16:0; C18:0; C18:1; C18:2	4

**Supplementary Table 5** List of detected metabolites in oxylipins analysis.

<b>Metabolite</b>	<b>Systematic name</b>	<b>Formula</b>	<b>Lipid Maps ID</b>
<b>10-HDoHE</b>	(+/-)-10-hydroxy-4Z,7Z,11E,13Z,16Z,19Z-docosahexaenoic acid	C22H32O3	LMFA04000027
<b>11,12-DiHETrE</b>	(±)11,12-dihydroxy-5Z,8Z,14Z-eicosatrienoic acid	C20H34O4	LMFA03050008
<b>11-HDoHE</b>	(+/-)-11-hydroxy-4Z,7Z,9E,13Z,16Z,19Z-docosahexaenoic acid	C22H32O3	LMFA04000028
<b>11-HETE</b>	11R-hydroxy-5Z,8Z,12E,14Z-eicosatetraenoic acid	C20H32O3	LMFA03060028
<b>12,13-DiHOME</b>	(+/-)-12,13-dihydroxy-9Z,15Z-octadecadienoic acid	C18H32O4	LMFA02000046
<b>12,13-EpOME</b>	(+/-)-12(13)-epoxy-9Z-octadecenoic acid	C18H32O3	LMFA02000038
<b>12-HETE</b>	12-hydroxy-5Z,8Z,10E,14Z-eicosatetraenoic acid	C20H32O3	LMFA03060088
<b>12S-HHTrE</b>	12S-hydroxy-5Z,8E,10E-heptadecatrienoic acid	C17H28O3	LMFA03050002
<b>13,14-dihydro-PGF2a</b>	9S,11R,15S-trihydroxy-5Z-prostenoic acid	C20H36O5	LMFA03010079
<b>13-HDoHE</b>	(+/-)-13-hydroxy-4Z,7Z,10Z,14E,16Z,19Z-docosahexaenoic acid	C22H32O3	LMFA04000029
<b>13-HODE</b>	13S-hydroxy-9Z,11E-octadecadienoic acid	C18H32O3	LMFA01050349
<b>14,15-DiHETrE</b>	14,15-dihydroxy-5Z,8Z,11Z-eicosatrienoic acid	C20H34O4	LMFA03050010
<b>14-HDoHE</b>	(+/-)-14-hydroxy-4Z,7Z,10Z,12E,16Z,19Z-docosahexaenoic acid	C22H32O3	LMFA04000030
<b>15S-HETrE</b>	15S-hydroxy-8Z,11Z,13E-eicosatrienoic acid	C20H34O3	LMFA03050007
<b>17,18-DiHETE</b>	(+/-)-17,18-dihydroxy-5Z,8Z,11Z,14Z-eicosatetraenoic acid	C20H32O4	LMFA03060078
<b>19,20-DiHDPA</b>	(±)19,20-dihydroxy-4Z,7Z,10Z,13Z,16Z-docosapentaenoic acid	C22H34O4	LMFA04000043
<b>20-carboxy-LTB4</b>	5S,12R-dihydroxy-6Z,8E,10E,14Z-eicosatetraene-1,20-dioic acid	C20H30O6	LMFA03020016
<b>5,6-DiHETrE</b>	5S,6S-dihydroxy-7E,9E,11Z,14Z-eicosatetraenoic acid	C20H32O4	LMFA03060018
<b>5-HETE</b>	5S-hydroxy-6E,8Z,11Z,14Z-eicosatetraenoic acid	C20H32O3	LMFA03060002
<b>8,9-DiHETrE</b>	8,9-dihydroxy-5Z,11Z,14Z-eicosatrienoic acid	C20H34O4	LMFA03050006
<b>8-HETE</b>	(±)8-hydroxy-5Z,9E,11Z,14Z-eicosatetraenoic acid	C20H32O3	LMFA03060086
<b>9,10-DiHOME</b>	9,10-dihydroxy-12Z-octadecenoic acid	C18H34O4	LMFA01050350

---

<b>9,10-EpOME</b>	(+/-)-9(10)-epoxy-12Z-octadecenoic acid	C18H32O3	LMFA02000037
<b>9,12,13-TriHOME</b>	9S,12S,13S-trihydroxy-10E-octadecenoic acid	C18H34O5	LMFA02000014
<b>9-HODE</b>	9S-hydroxy-10E,12Z-octadecadienoic acid	C18H32O3	LMFA01050278
<b>9-HOTrE</b>	9S-hydroxy-10E,12Z,15Z-octadecatrienoic acid	C18H30O3	LMFA02000024
<b>LTB4</b>	5S,12R-dihydroxy-6Z,8E,10E,14Z-eicosatetraenoic acid	C20H32O4	LMFA03020001
<b>PGE2</b>	9-oxo-11R,15S-dihydroxy-5Z,13E-prostadienoic acid; Prostin E2	C20H32O5	LMFA03010003
<b>TXB1</b>	9S,11,15S-trihydroxy-thrombox-13E-enoic acid	C20H36O6	LMFA03030008
<b>TXB2</b>	9S,11,15S-trihydroxy-thromboxa-5Z,13E-dien-1-oic acid	C20H34O6	LMFA03030002
<b>TXB3</b>	9S,11,15S-trihydroxy-thromboxa-5Z,13E,17Z-trien-1-oic acid	C20H32O6	LMFA03030006

---

**Supplementary Table 6:** List of detected metabolites in amines analysis

<b>Metabolite</b>	<b>Chemical formula</b>	<b>HMDB ID</b>	<b>InChI Key</b>
<b>1-Methylhistidine</b>	C7H11N3O2	HMDB00001	BRMWTNUJHUMWMS-LURJTMIESA-N
<b>3-Aminoisobutyric acid</b>	C4H9NO2	HMDB00452	QWCKQJZIFLGMSD-VKHKMYHEASA-N
<b>4-Hydroxyproline</b>	C5H9NO3	HMDB00725	PMMYEEVYMWASQN-DMTCNVIQSA-N
<b>Alanine</b>	C3H7NO2	HMDB00161	QNAYBMKLOCPYGJ-REOHCLBHSA-N
<b><math>\alpha</math>-aminobutyric acid</b>	C4H9NO2	HMDB03911	QCHPKSFMHDPSNR-UHFFFAOYSA-N
<b>Arginine</b>	C6H14N4O2	HMDB00517	ODKSFYDXFIFQN-BYPYZUCNSA-N
<b>Asparagine</b>	C4H8N2O3	HMDB00168	DCXYFEDJOCNFAF-REOHCLBHSA-N
<b>Aspartic acid</b>	C4H7NO4	HMDB00191	CKLJMWTZIZZHCS-REOHCLBHSA-N
<b>Citrulline</b>	C6H13N3O3	HMDB00904	RHGKLRLOHDJJDR-BYPYZUCNSA-N
<b>Cystathionine</b>	C7H14N2O4S	HMDB00099	ILRYLPWNYFXEMH-WHFBIAKZSA-N
<b>Cysteine</b>	C3H7NO2S	HMDB00574	XUJNEKJLAYXESH-REOHCLBHSA-N
<b>Ethanolamine</b>	C2H7NO	HMDB00149	HZAXFHJVJLSVMW-UHFFFAOYSA-N
<b>Glutamic acid</b>	C5H9NO4	HMDB00148	WHUUTDBJXRKMK-VKHKMYHEASA-N
<b>Glutamine</b>	C5H10N2O3	HMDB00641	ZDXPYRJPNDTMRX-VKHKMYHEASA-N
<b>Glutathione</b>	C10H17N3O6S	HMDB00125	RWSXRVCMGQZWBV-WDSKDSINSA-N
<b>Glycine</b>	C2H5NO2	HMDB00123	DHMQDGOQFOQNFH-UHFFFAOYSA-N
<b>Glycylglycine</b>	C4H8N2O3	HMDB11733	YMAWOPBAYDPSLA-UHFFFAOYSA-N
<b>Histidine</b>	C5H9N3	HMDB00870	NTYJJOPFIAHURM-UHFFFAOYSA-N
<b>Homocysteine</b>	C4H9NO2S	HMDB00742	FFFHZYDWPBMMWHY-UHFFFAOYSA-N
<b>Isoleucine</b>	C6H13NO2	HMDB00172	AGPKZVBTJJNPAG-WHFBIAKZSA-N
<b>Kynurenine</b>	C10H12N2O3	HMDB00684	YGPSJZOEDVAXAB-QMMMGPBSA-N
<b>Leucine</b>	C6H13NO2	HMDB00687	ROHFNLRQFUQHCH-YFKPBYRVSA-N
<b>Lysine</b>	C6H14N2O2	HMDB00182	KDXKERNBIXSRK-YFKPBYRVSA-N
<b>Methionine</b>	C5H11NO2S	HMDB00696	FFEARJCKVFRZRR-BYPYZUCNSA-N
<b>Methionine sulfoxide</b>	C5H11NO3S	HMDB02005	QEFRNWWLZKMPFJ-UHFFFAOYSA-N
<b>O-Phosphoethanolamine</b>	C2H8NO4P	HMDB00224	SUHOOTKUPISOBE-UHFFFAOYSA-N
<b>Ornithine</b>	C5H12N2O2	HMDB00214	AHLPHDHHMVZTML-BYPYZUCNSA-N

---

<b>Phenylalanine</b>	C9H11NO2	HMDB00159	COLNVLDHVKWLRT-QMMMGPBSA-N
<b>Pipecolic acid</b>	C6H11NO2	HMDB00716	HXEACLLIILLPRG-YFKPBYSVA-N
<b>Proline</b>	C5H9NO2	HMDB00162	ONIBWKKTOPOVIA-BYPYZUCNSA-N
<b>Sarcosine</b>	C3H7NO2	HMDB00271	FSYKKLYZXJSPZ-UHFFFAOYSA-N
<b>Serine</b>	C3H7NO3	HMDB00187	MTCFGRXMJLQNBG-REOHCLBNSA-N
<b>Serotonin</b>	C10H12N2O	HMDB00259	QZAYGJVTTNCVMB-UHFFFAOYSA-N
<b>s-Methylcysteine</b>	C4H9NO2S	HMDB02108	IDIDJDIHTAOVLG-VKHMHEASA-N
<b>Taurine</b>	C2H7NO3S	HMDB00251	XOAAWQZATWQOTB-UHFFFAOYSA-N
<b>Threonine</b>	C4H9NO3	HMDB00167	AYFVYJQAPQTCCC-GBXISLNSA-N
<b>Tryptophan</b>	C11H12N2O2	HMDB00929	QIVBCDIJAJPQS-VIFPVBQESA-N
<b>Tyrosine</b>	C9H11NO3	HMDB00158	OUYCCASQSFEME-QMMMGPBSA-N
<b>Valine</b>	C5H11NO2	HMDB00883	KZSNJWFQEVHDMF-BYPYZUCNSA-N
<b>γ-glutamylalanine</b>	C10H17N3O6	HMDB11738	JBFYFLXEJFQWMU-WDSKDSNSA-N

---

**Supplementary Table 7.** Classification table of predicted good- and non-responders and observed good- and non-responders.

		Observed	Good resp	Non-resp	Sensitivity	Specificity	PPV	NPV	MR
		Predicted							
Clinical model	Good response	43	20	0.782	0.600	0.683	0.714	0.305	
	Non-response	12	30						
Combined model	Good response	39	5	0.709	0.900	0.887	0.738	0.200	
	Non-response	16	45						

PPV, positive predictive values; NPV, negative predictive values; MR, misclassification rate.

**Supplementary Table 8:** Net reclassification index of prediction models for sensitivity analysis.

<b>Observed response</b>	<b>Predicted by clinical model</b>	<b>Predicted by combined model</b>	
		<b>Non-response</b>	<b>Response</b>
<b>Non-responders (n=50)</b>	<b>Non-response</b>	30 (equal)	0 (worsening)
	<b>Response</b>	15 (improvement)	5 (equal)
<b>Responders (n=100)</b>	<b>Non-response</b>	23 (equal)	12 (improvement)
	<b>Response</b>	20 (worsening)	45 (equal)

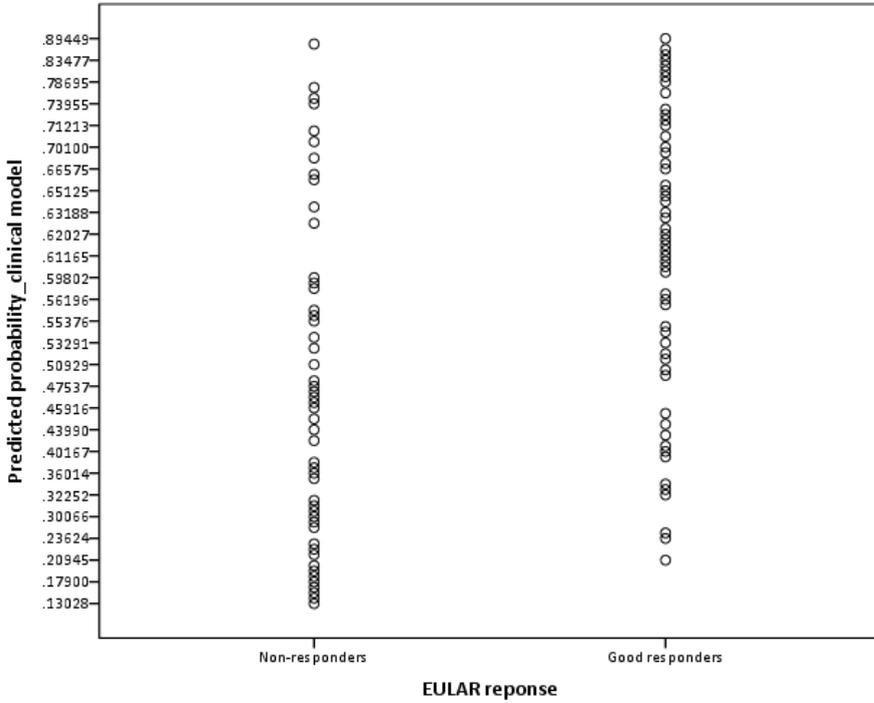
**Supplementary Table 9:** Metabolites cross-sectionally associated with either baseline DAS28, ESR or CRP ( $p < 0.05$ ) based on the complete cohort of bDMARD users ( $n=231$ ). The log-transformed and autoscaled metabolites were put in a linear GEE on DAS28, ESR and CRP separately, together with possible influencing clinical parameters (age, gender, menopausal status, BMI, smoking status, alcohol consumption, concomitant csDMARDs (MTX, HCQ, LEF, GCs) and non-DMARDs (statins, antihypertensive drugs, bisphosphonates and NSAIDs). Shown are the regression coefficients and  $p$ -values per metabolite, only of those metabolites that associated to either DAS28, ESR or CRP.

	LnCRP		DAS28		lnESR	
	Coeff	$p$ -value	Coeff.	$p$ -value	Coeff	$p$ -value
<i>sn2</i> -LPC (14:0)	-.666	.004	-.066	.323	-.604	.002
<i>sn1</i> -LPC (14:0)	-.713	.000	-.080	.255	-.606	.000
<i>sn2</i> -LPC (16:1)	-.525	.194	-.122	.098	-.759	.000
<i>sn1</i> -LPC (18:3-w3w6)	-.223	.487	.048	.564	-.646	.000
<i>sn1</i> -LPC (20:5)	-1.023	.045	-.007	.944	-.556	.138
<i>sn1</i> -LPC (16:1)	-.633	.099	-.086	.213	-.744	.000
<i>sn2</i> -LPC (18:2)	-.396	.078	-.785	.001	-.468	.022
<i>sn1</i> -LPC (15:0)	-.960	.002	-1.176	0.243	-.578	.002
<i>sn1</i> -LPC (18:2)	-.397	.134	-.753	.001	-.437	.031
<i>sn1</i> -LPC (20:4)	-.418	.049	-.056	.475	-.357	.029
<i>sn2</i> -LPC (20:3)-w3w6	-.093	.510	-.071	.280	-.442	.028
LPE (22:5-w3)	-.561	.038	-.071	.450	-.300	.058
<i>sn1</i> -LPC (22:5-w3)	-.710	.009	-.109	.167	-.573	.003
<i>sn1</i> -LPC (20:3-w3w6)	-.268	.081	-.090	.200	-.513	.006
LPE (22:5-w6)	-.437	.223	-.730	.012	-.602	.002
<i>sn1</i> -LPC (22:5-w6)	-.379	.004	-.725	.000	-.596	.000
<i>sn1</i> -LPC (18:1)	-.355	.182	-.081	.314	-.439	.032
<i>sn1</i> -LPC (20:3-w9)	.000	.999	-.132	.059	-.482	.004
<i>sn1</i> -LPC (22:4)	-.488	.005	-.049	.517	-.677	.000
LPC (O-16:0)	-.289	.015	-.729	.000	-.289	.059
LPC (O-18:2)	-.441	.326	-.103	.155	-.778	.005
<i>sn1</i> -LPC (20:2)	-.351	.077	-.011	.875	-.506	.010
<i>sn1</i> -LPC (19:0)	-.297	.517	-.662	.004	-.276	.147
LPC (O-18:0)	-.537	.006	-.057	.520	-.212	.177
FA (14:1)	.627	.000	.607	.000	.444	.000
FA (14:0)	.278	.037	.059	.424	.407	.000
FA (20:5-w3)	.025	.872	.009	.918	.496	.000
FA (18:3-w3w6)	.402	.007	.090	.263	.371	.001
FA (16:1)	.441	.006	.640	.000	.477	.000
FA (15:0)	.232	.034	.036	.633	.290	.003
FA (22:6)	.434	.169	-.012	.879	.726	.000
FA (20:4-w6)	.095	.574	.013	.846	.482	.000
FA (18:2)	.443	.039	.067	.314	.459	.000
FA (22:5-w3)	.363	.111	.590	.000	.489	.000
FA (17:1)	.475	.003	.052	.540	.450	.000
FA (22:5-w6)	.380	.198	-.008	.917	.571	.000
FA (20:3-w3w6)	.194	.206	-.012	.869	.502	.000
FA (16:0)	.244	.023	.421	.000	.331	.000
FA (20:3-w9)	.226	.013	.075	.814	-.045	.749
FA (18:1)	.547	.009	.583	.000	.471	.000
FA (22:4)	.438	.060	.597	.000	.460	.000

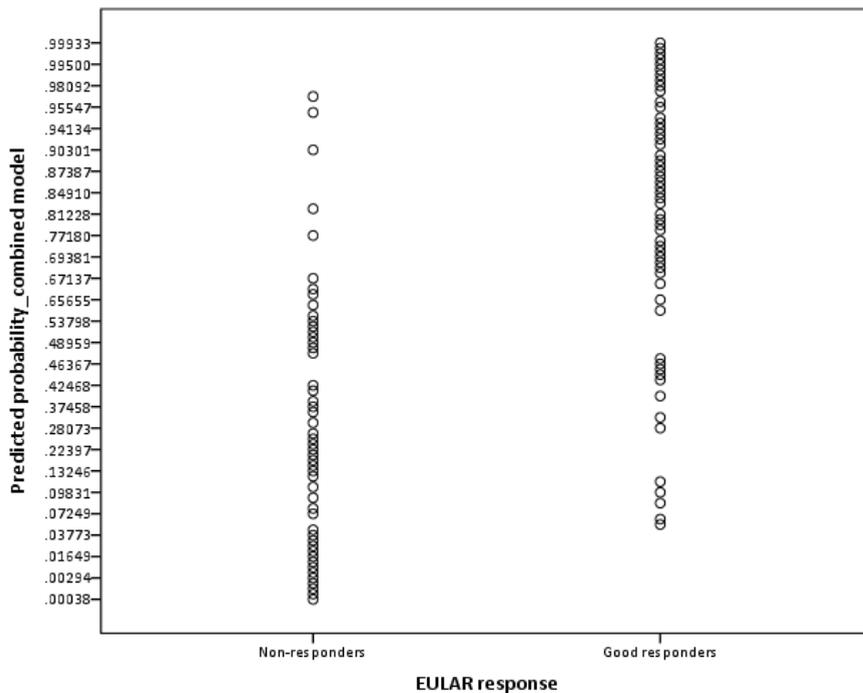
FA (17:0)	.220	.001	.353	.001	.215	.013
FA (20:2)	1.032	0.064	.002	.989	.406	.002
FA (18:0)	.175	.015	.351	.000	.223	.016
FA (20:1)	.451	.039	.093	.220	.427	.000
FA (24:1)	.349	.151	.529	.000	.369	.004
TXB3	.486	.000	.396	.183	.095	.436
TXB1	.404	.002	.937	.000	.596	.001
TXB2	.467	.002	.502	.213	.014	.931
9,12,13-TriHOME	-.081	.714	-.646	.073	-.505	.008
PGE2	.392	.015	.736	.021	.362	.017
13,14-dihydro-PGF2a	-.352	.008	-.769	.000	-.411	.060
19,20-DiHDPA	-.683	0.020	-.036	.924	.006	.984
14,15-DiHETrE	-.746	.000	.368	.190	.234	.457
11,12-DiHETrE	-.586	.006	.475	.051	.316	.205
9-HOTrE	.068	.686	.359	.020	.292	.019
8,9-DiHETrE	-.611	.005	-.216	.427	.070	.669
5,6-DiHETrE	-.997	.005	.255	.644	.054	.900
9-HODE	.012	.933	.255	.183	.260	.010
14-HDoHE	.427	.136	.749	.041	.548	.003
10-HDoHE	.238	.295	.675	.009	.645	.000
13-HDoHE	.270	.362	.648	.086	.580	.003
11-HETE	.396	.001	.895	.000	.420	.005
11-HDoHE	.242	.063	.860	.000	.526	.000
12-HETE	.463	.003	.708	.048	.412	.006
8-HETE	.243	.020	.654	.000	.413	.001
15S-HETrE	.281	.013	.730	.000	.392	.012
5-HETE	.273	.010	.729	.000	.498	.000
Ornithine	.483	.003	.061	.832	.024	.905
Sarcosine	-.386	.065	-.664	.003	-.558	.000
Alpha-aminobutyric acid	.181	.446	-.433	.035	-.166	.479
Proline	.284	.028	-.051	.573	.103	.373
Valine	.840	.000	-.017	.829	.268	.134
Threonine	.106	.678	-.205	.006	-.102	.655
Cysteine	.232	.107	.085	.392	.280	.017
Taurine	.626	.279	.124	.086	.461	.005
Isoleucine	.638	.001	.041	.606	.197	.218
Leucine	.642	.000	.323	.160	.190	.284
4-Hydroxyproline	.285	.024	-.007	.952	-.078	.425
Asparagine	-.033	.938	-.150	.024	-.146	.553
Glycylglycine	-.030	.944	-.164	.015	-.248	.324
Histidine	.051	.904	-.284	.000	-.113	.576
Arginine	.434	.012	-.059	.389	.211	.170
Serotonin	.977	.000	-.044	.655	.684	.000
Tryptophan	.791	.000	-.113	.216	-.108	.770
Kynurenine	.529	.026	.082	.338	-.004	.983
Gamma-glutamylalanine	.941	.000	-.111	.099	.209	.413
Glutathione	.388	.016	.433	.006	.472	.000
sn2-LPC (20:5)	-.728	.136	-.042	.671	-.435	.274
sn2-LPC (18:3-w3w6)	.138	.521	.147	.622	-.314	.073
LPE (20:5)	-.237	.409	.226	.459	.015	.947
sn2-LPC (20:4)	-.259	.127	-.064	.413	-.321	.058
sn2-LPC (22:6)	-.444	.085	-.081	.300	.088	.677

<i>sn1</i> -LPC (22:6)	-.456	.095	-.069	.369	.147	.561
LPE (18:2)	-.412	.338	-.026	.788	-.075	.770
LPE (22:6)	-.163	.429	.359	.135	.346	.070
LPE (20:4)	-.484	.177	-.008	.919	.070	.755
<i>sn2</i> -LPC (16:0)	-.124	.592	-.015	.852	-.118	.428
<i>sn2</i> -LPC (18:1)	-.110	.561	-.057	.451	-.298	.126
<i>sn1</i> -LPC (16:0)	-.191	.260	-.037	.626	-.180	.195
LPE (16:0)	-.283	.174	.021	.814	-.109	.401
LPE (20:3-w3w6)	-.359	.336	-.052	.550	-.023	.900
LPE (18:1)	-.316	.394	-.037	.747	.178	.439
<i>sn2</i> -LPC (17:0)	.133	.449	.203	.492	-.194	.199
<i>sn2</i> -LPC (18:0)	-.085	.550	-.048	.504	-.086	.464
LPC (O-18:1)	-.179	.174	-.022	.757	-.142	.198
LPE (18:0)	-.130	.396	.056	.547	.247	.125
<i>sn1</i> -LPC (18:0)	-.077	.542	-.051	.476	-.064	.556
<i>sn1</i> -LPC (20:1)	-.164	.356	.021	.713	-.036	.704
FA (20:0)	.301	.110	.035	.630	.191	.151
FA (24:0)	.184	.334	.076	.404	.144	.347
20-carboxy-LTB4	.102	.549	.262	.212	.130	.313
17,18-DiHETE	-.405	.053	.549	.081	.153	.624
LTB4	-.056	.691	.185	.550	.228	.188
12,13-DiHOME	.048	.823	.244	.205	.092	.479
9,10-DiHOME	-.108	.623	.133	.543	.089	.529
12S-HHTrE	.209	.592	.528	.189	-.172	.414
13-HODE	-.111	.446	.065	.774	.160	.141
12,13-EpOME	-.066	.665	.124	.476	.169	.065
9,10-EpOME	-.090	.554	.111	.536	.199	.059
Ethanolamine	.091	.878	.039	.620	.400	.102
Lysine	.191	.291	-.072	.246	.021	.915
Glycine	.009	.987	-.037	.673	-.230	.453
Alanine	.138	.602	-.097	.144	-.080	.722
3-Aminoisobutyric acid	-.052	.872	-.132	.095	-.308	.187
Serine	.084	.768	-.120	.152	-.265	.250
Cystathionine	-.178	.221	-.095	.355	-.233	.065
Pipecolic acid	.315	.070	-.183	.126	-.009	.963
Aspartic acid	.180	.562	.173	.585	.109	.577
s-Methylcysteine	.005	.988	.066	.528	.243	.327
Homocysteine	-.063	.839	.145	.072	.340	.498
O-Phosphoethanolamine	-.127	.493	.052	.588	.097	.430
Glutamine	.109	.751	-.049	.500	.228	.226
Glutamic acid	-.248	.413	.014	.861	-.304	.111
Methionine	.048	.844	-.052	.413	-.061	.735
Methionine sulfoxide	.651	.055	.040	.595	.283	.204
Phenylalanine	.242	.298	.052	.831	-.012	.951
1-Methylhistidine	.528	.113	.101	.123	.284	.127
Citrulline	-.309	.566	.285	.568	-.392	.326
Tyrosine	-.091	.763	-.014	.852	.157	.398

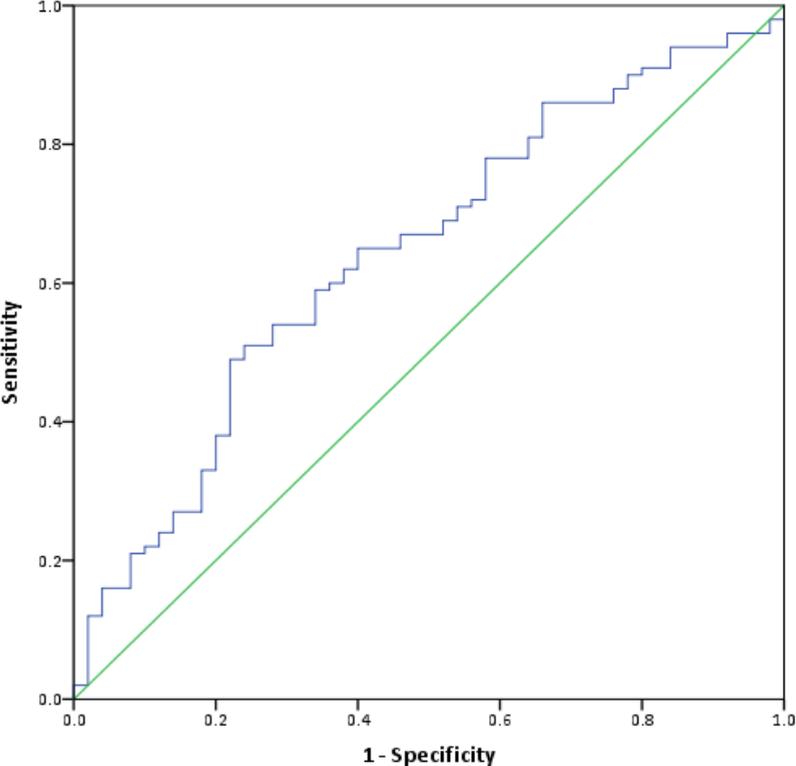
**Supplementary Figure 1:** Scatter plot of the clinical model between good responders and non-responders to TNFi therapy. Patients were grouped according to their observed responses on the x-axis; whereas the y-axis represents the predicted probability calculated by the clinical model. The pseudo R-square, as a measure for the explained variance in the observed response by the model was 0.147 (Cox & Snell).



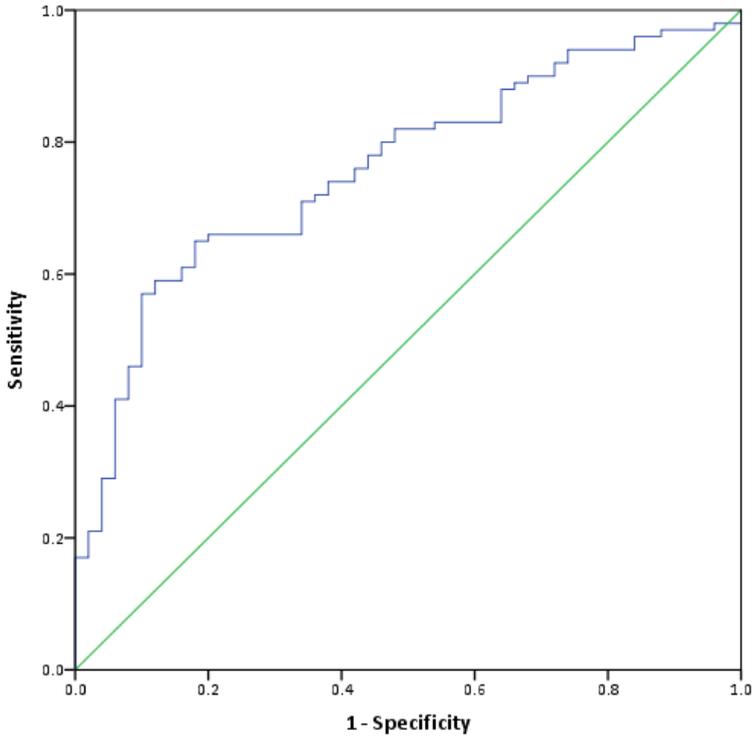
**Supplementary Figure 2:** Scatter plot of the combined model between good responders and non-responders to TNFi therapy. Patients were grouped according to their observed responses on the x-axis; whereas the y-axis represents the predicted probability calculated by the combined model. The pseudo R-square, as a measure for the explained variance in the observed response by the model was 0.433 (COX & Snell)



**Supplementary Figure 3.** ROC curve for the clinical model containing non-, moderate- and good responders to TNFi therapy. The AUC-ROC was 0.641 (95% CI: 0.548-0.734).



**Supplementary Figure 4:** ROC curve for the combined model non-, moderate- and good responders to TNFi therapy. The AUC-ROC was 0.760 (95% CI: 0.682-0.837)





## Chapter 9

### **Differences between serum polar lipid profiles of male and female rheumatoid arthritis patients in response to glucocorticoid treatment**

Bart V.J. Cuppen<sup>#</sup>, Junzeng Fu<sup>#</sup>, Paco M.J. Welsing, Herman van Wietmarschen, Amy C. Harms, Ruud Berger, Slavik Koval, Ruth D.E. Fritsch-Stork Johannes W.J. Bijlsma, Thomas Hankemeier, Jan van der Greef, and Floris P.J.G. Lafeber

<sup>#</sup> These authors contributed equally

## ABSTRACT

### Objective

As there are pharmacological differences between males and females, and glucocorticoid (GC) treatment is associated with increased cardiovascular mortality rate in rheumatoid arthritis (RA) patients, it is important to study serum polar lipid profiles of male and female patients in response to GC therapy. Gender differences may require an adjustment to the treatment strategy for a selection of patients.

### Methods

Serum samples from 281 RA patients were analysed using a targeted lipidomics platform. The differences in GC use and gender on polar lipid profiles were cross-sectionally examined by multiple linear regression, while correcting for confounding factors.

### Results

Differences in polar lipids between GC-users and non-GC-users in females and males were merely restricted to lysophospholipids (lysophosphatidylcholines and lysophosphatidylethanolamines). Lysophospholipids in female patients treated with GCs were significantly higher than female patients not treated with GCs ( $p=6.0 \times 10^{-6}$ ), whereas no significant difference was observed in male GC users *versus* non-users ( $p=0.397$ ).

### Conclusion

The lysophospholipid profiles in response to GCs were significantly different between male and female RA patients, which may have implications for the cardiovascular risk of GC treatment.

## INTRODUCTION

Rheumatoid arthritis (RA) is an auto-immune disease with unresolved aetiology, which predominantly occurs in females [1]. Glucocorticoids (GCs) have been prescribed for the treatment of RA for decades, and are considered to be effective drugs in reducing inflammation and preventing joint destruction [2]. The role of gender as a crucial factor in drug studies is becoming increasingly appreciated [3]. Several studies have investigated the effects of gender on clinical pharmacology for GCs on healthy volunteers, and showed gender specific differences in GC pharmacokinetics and pharmacodynamics [4, 5]. However, these differences are balanced by complementary GC clearance and GC sensitivity. Therefore, these gender differences do not necessitate GC dose adjustments in clinical practice.

Glucocorticoids are known to undesirably affect triacylglyceride and fatty acid metabolism [6]. It is thus conceivable that GC-induced changes of lipid profiles in RA patients also show gender dependence, apart from the gender differences in lipid metabolism seen in the general population [7]. Because GC treatment is associated with increased cardiovascular mortality rate in RA patients [8, 9], it is important to study these gender-based lipid differences as they may require an adjustment in treatment strategy for a selection of the patients.

In the present study, we measured circulating polar lipids, such as lysophospholipids and free fatty acids, in the serum of RA patients using a targeted lipidomics platform, because polar lipids are crucial intermediates in lipid metabolism. Lipid profiles between GC users and non-GC users were examined and analysed for gender differences. Our results suggested that the lipid profile is more affected by GC treatment in female RA patients. In particular the levels of lysophospholipids were more elevated in female users compared to non-users. In males the differences of lysophospholipid levels in GC users were not significant compared to non-users.

## METHODS

Subjects in this study were participants of an observational study BiOCURA, in which RA patients initiating or switching from biological therapy were recruited [10]. Blood samples

were collected before initiating the biological therapy and immediately processed into serum. Serum samples were stored at -80°C until use for lipidomic analyses. At the time of sampling, 42.35 % of patients were receiving GCs, including prednisolone and prednisone, at varying dosages (all medications are listed in Supplementary Table 1). The study was approved by the medical ethics committee of the University Medical Center Utrecht and the institutional review boards of the participating centres (see Acknowledgements). Written informed consent was obtained from each patient.

The operating procedures of the targeted lipidomics platform are optimised from the previously published method [11]. Polar lipids are extracted by methanol from serum samples, and analysed by liquid chromatography-mass spectrometry, covering the low abundance lipid species, including free fatty acids and lysophospholipids—lysophosphatidylcholines (LPCs) and lysophosphatidylethanolamines (LPEs). Details of the procedures are described in the Supplementary Method.

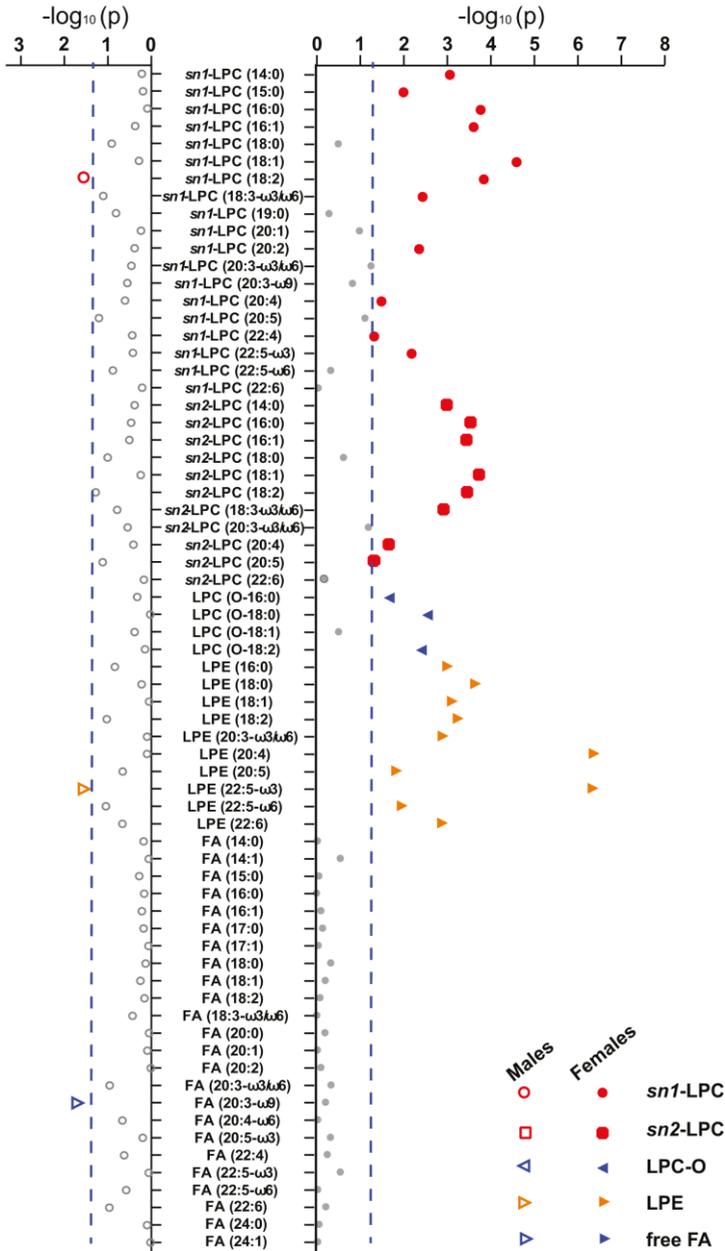
A schematic overview of the statistical analyses is provided in Supplementary Figure 1. Initially, the obtained lipid data set was log transformed and standardised into Z-scores to produce normalised and auto-scaled data (mean=0, SD=1). Then, the differences in lipids between GC users and non-GC-users were calculated and tested for significance by independent t-tests for male and female subjects separately. In parallel, principal component analysis (PCA) was performed on all detected lipids to elucidate the correlation structure of the metabolites. By combining the results of the t-tests, PCA, and prior biological knowledge, a decision was made on which lipids can be clustered into a new lipid score, in order to have one overall outcome for subsequent analyses. For each patient, the score was computed by summing the standardized values of lipids and dividing this by the number of included lipids ( $\sum_i^N \text{standardised lipid}_i / N$ , with N=number of lipids clustered). This value represents both the average of lipids, as well as the patients' relative deviation from the mean lipid score in standard deviations (SDs). Multiple linear regression analysis was then conducted to study the effect of GC use between males and females on the lipid score. We entered the following subgroups in the model—female-GC user (n=77), female-GC non-user (n=136), male-GC user (n=42), and male-GC non-user (n=26) together with the clinical parameters (listed in Supplementary Table 1) as a full model, while setting the subgroup “Female-GC non-user” as the reference group (i.e. the

intercept of the regression model). To arrive at a final model, backward elimination was applied on the full model by excluding clinical parameters one by one on p-values (starting from highest to lowest p-value). Parameters were excluded, only when the change in the regression coefficients after exclusion was < 10% for all four subgroups; otherwise the clinical parameter was kept in the model as a confounder. In order to explore the difference in lipid score between GC users and non-users in males, the reference group in the final model was switched to “male GC non-users”.

## RESULTS AND DISCUSSION

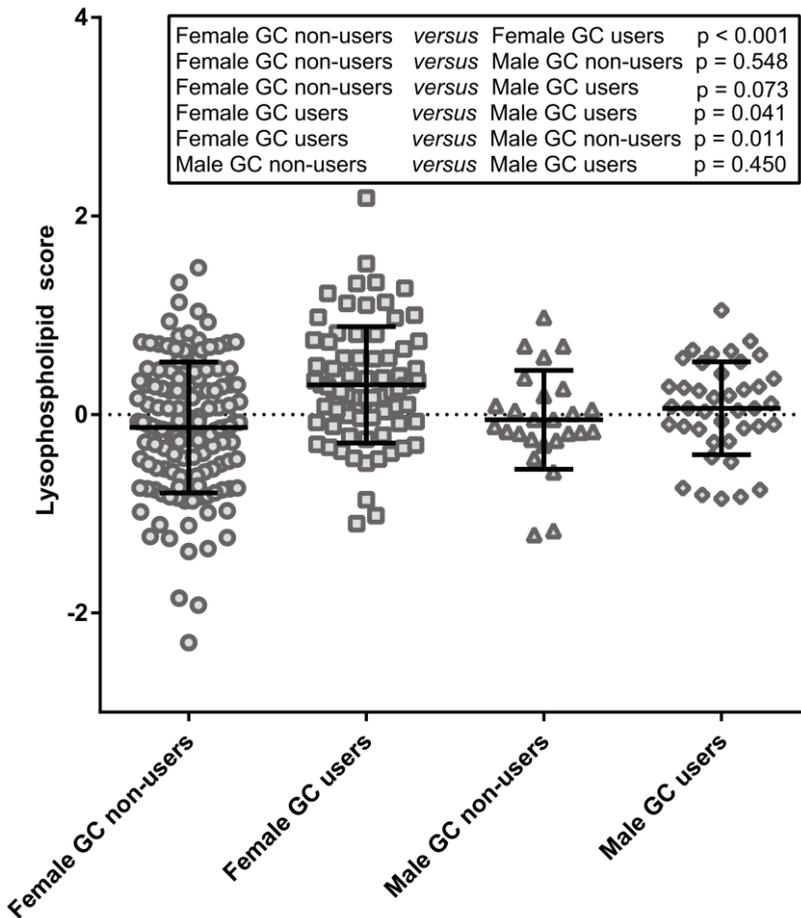
In the patient cohort (n=281), there were more males taking GCs than females (61.8% vs. 35.8%,  $p < 0.01$ ). There were no significant differences in disease activity among the relevant groups (Supplementary Table 1). However, the number of smokers was especially high in males ( $p < 0.001$ ), positive rheumatoid factor was high amongst male GC users ( $p = 0.004$ ), GC users used less non-steroidal anti-inflammatory drugs (NSAIDs) ( $p = 0.010$ ) and more bisphosphonates (to reduce the risk of GC-induced osteoporosis) ( $p < 0.001$ ).

By applying an established lipidomics platform, serum lipid profiles of the 281 RA patients were analysed, covering 44 lysophospholipids and 24 free fatty acids (Supplementary Table 2). T-tests between GC users and non-users were performed (Supplementary Table 3; a graphical representation is shown in Figure 1). In females, we identified 10 LPEs and 22 LPCs which were significantly higher in GC users than in GC non-users ( $p < 0.05$ ), whereas no differences were found in fatty acids. However, in male subjects only one LPE (LPE (20:5),  $p = 0.029$ ) and LPC (*sn1*-LPC (18:2),  $p = 0.027$ ), which were significant in females, were found significantly higher in GC users whereas one fatty acid (FA (20:3- $\omega$ 9),  $p = 0.021$ ) was significant lower in male GC users among all measured FAs. These results suggest that the GCs have a more pronounced impact on female lysophospholipid profile compared to males.



**Figure 1:** Graphical representation of t-tests results in lipid clustering. Independent sample t-tests were performed on all 68 metabolites on glucocorticoid (GC) users versus non-users, for both genders separately. The x-axis shows the logarithm of p-values per metabolite for males and females, the blue dash line represents p-value of 0.05.

The parallel PCA analysis on 68 metabolites showed that the loading scores of all 32 significant lysophospholipids were larger than 0.4 in the first component, thus highly correlated with each other (Supplementary Figure 2, Supplementary Table 4). A new score representing the lysophospholipid levels could therefore be computed, by calculation of the mean of all significant lysophospholipids. As shown in Figure 2, the absolute lysophospholipid scores were significantly different between female GC users and non-



**Figure 2:** Lysophospholipid score in male patients without ( $n = 26$ ) or with glucocorticoid (GC) treatment ( $n = 42$ ), and in female patients without ( $n=136$ ) or with GC treatment ( $n = 77$ ). Horizontal bars indicate mean values and standard deviation. One-way ANOVA with Fisher's LSD was used to compare the means of score among subgroups.

users ( $p < 0.001$ ), whereas no difference was seen in males ( $p = 0.450$ ). In addition, female GC users showed significantly higher values compared to male GC users ( $p = 0.041$ ), whereas no difference was seen between female and male non-users ( $p = 0.548$ ).

After backward elimination for potential confounders, a final model for the lysophospholipid score was established (Table 1). Subgroups of males and female GC users all had significant positive coefficients, which suggest that the lysophospholipid score of these subgroups was significantly higher compared to females not using GCs. In particular, females using GCs had a significant increase in mean lysophospholipid score of 0.398 ( $p = 6 \times 10^{-6}$ ), compared to females not using GCs. Interestingly, the difference between male and female non-users was not significant before correcting for confounders ( $p = 0.548$ , Figure 2), while it was significant in the final model (coefficient = 0.310,  $p = 0.041$ ). The difference between GC users and non-users in males was non-significant (0.122,  $p = 0.397$ ; model with “male GC non-users” as a reference shown in Supplementary Table 5).

**Table 1:** Final regression model investigating the association between gender and glucocorticoid (GC) use on the lysophospholipid score corrected for confounders. Shown is the difference in mean lysophospholipid score for subgroups compared to females not using GCs (as reference group).

Variables	Coefficients <sup>a</sup> (95%-CI)	p value
<b>Female GC non-user (reference group)</b>	-0.161 (-0.729 to 0.408)	0.580
<b>Female GC user</b>	0.398 (0.229 to 0.567)	<b>6.0 E-6</b>
<b>Male GC non-user</b>	0.310 (0.015 to 0.604)	<b>0.041</b>
<b>Male GC user</b>	0.432 (0.164 to 0.700)	<b>1.7 E-3</b>

<sup>a</sup> Coefficients indicate the changes in mean lysophospholipid score, adjusting for: Age, BMI, menopausal status, RF positive, log-transformed CRP, log transformed ESR, DAS28, and concomitant drugs (methotrexate, hydroxychloroquine, anti-diabetic drug and NSAIDs). BMI, body mass index; CI, confidence interval; CRP, C-reactive protein; GC, glucocorticoid; RF, rheumatoid factor; ACPA, anti citrullinated protein antibody; ESR, erythrocyte sedimentation rate; DAS28, disease activity score based on a 28-joint count; NSAIDs, non-steroid anti-inflammatory drugs

Therefore, after correcting for confounders, the lysophospholipids score of female RA patients taking GCs was still significantly higher than the female patients not taking GCs, whereas in male RA patients no difference in lysophospholipid score was seen. This was consistent with the results of the t-tests on individual lipids and the uncorrected difference in lysophospholipid score.

Patients with RA already have a higher cardiovascular disease risk and this elevated risk is only partly explained by the increased prevalence of traditional cardiovascular risk factors such as age, gender, dyslipidaemia, hypertension, smoking, obesity and diabetes mellitus [12]. In addition, systemic inflammation and genetic factors also play a role [12]. More recently, GC use has been directly related to an (dose dependent) increase in cardiovascular death in RA [8]. However, in this study no effect on lipid profiles by different dosages was seen, as the factor low (<7.5mg) *versus* moderate to high (>7.5mg) dosage was excluded during confounder selection. A possible protective effect can be expected from concomitant use of hydroxychloroquine, which significantly lowered lysophospholipid scores in our study (decrease in mean lysophospholipid score=0.180, 95% CI (-0.347 to -0.013), p=0.035; Supplementary Table 5). It has also been reported to improve cholesterol levels, notably in those treated with GCs [13].

Lysophospholipids, including lysophosphatidylcholine (LPC) and lysophosphatidylethanolamine (LPE), are an abundant lipid species, mainly functioning as transporters for free fatty acids. The difference between LPC and LPE is based only on the functional head group— respectively choline or ethanolamine. The functions of LPEs are underreported, hampering their biological interpretation. Studies show that LPCs have properties resembling extracellular growth factors and signalling molecules [14]. *In vivo*, LPCs are generated from phospholipase A1/A2 catalysed hydrolysis of phosphatidylcholines, the basic component of membranes [15]. In addition, LPCs are released from phosphatidylcholines by the action of lecithin cholesterol acyltransferase in plasma [16]. Most of the circulating LPC is bound to albumin, but they are also a major component of lipoprotein particles, where they are a known constituent of oxidised low-density-lipoproteins (LDL) [17, 18], a well-known risk factor for cardiovascular diseases [19]. The lysophospholipid related gender differences are therefore potentially relevant

with respect to the risk of cardiovascular events in RA patients, which could eventually guide the adjustment of treatment strategies for either males or females.

As patients were included in BiOCURA based on the necessity of biological treatment and not GC treatment, it was only possible to use samples of users and non-users, but not before and after GC initiation. Future studies are needed to validate our results, preferably before and after initiation of GC treatment. Additionally the role of lipid profiles (including triglycerides, diglycerides, and sphingomyelins) in the association between GC use, gender and cardiovascular death, should be clarified, in order to fully understand and (specifically) prevent unwanted clinical (side-)effects.

## **CONCLUSION**

After correcting for confounding factors, lysophosphatidylcholines and lysophosphatidylethanolamines in female RA patients with GC treatment were significantly higher than in female patients not taking GCs, whereas in male RA patients these lysophospholipids levels were similar between GC users and non-users. These results could contribute to a better understanding and estimation of safety of GC drugs for male and female RA patients separately, particularly in relation to cardiovascular events.

## REFERENCES

1. Jutley G, Raza K, Buckley CD. New pathogenic insights into rheumatoid arthritis. 2015.
2. Hoes JN, Jacobs JWG, Boers M, Boumpas D, Buttgerit F, Caeyers N et al. EULAR evidence-based recommendations on the management of systemic glucocorticoid therapy in rheumatic diseases. *Annals of the rheumatic diseases* 2007;66: 1560-7.
3. Soldin OP, Mattison DR. Sex differences in pharmacokinetics and pharmacodynamics. *Clin Pharmacokinet* 2009;48 3: 143-57.
4. Lew KH, Ludwig EA, Milad MA, Donovan K, Middleton E, Ferry JJ et al. Gender-based effects on methylprednisolone pharmacokinetics and pharmacodynamics. *Clinical pharmacology and therapeutics* 1993;54: 402-14.
5. Magee MH, Blum RA, Lates CD, Jusko WJ. Prednisolone pharmacokinetics and pharmacodynamics in relation to sex and race. *Journal of clinical pharmacology* 2001;41: 1180-94.
6. Macfarlane DP, Forbes S, Walker BR. Glucocorticoids and fatty acid metabolism in humans: fuelling fat redistribution in the metabolic syndrome. *The Journal of endocrinology* 2008;197: 189-204.
7. Wang X, Magkos F, Mittendorfer B. Sex differences in lipid and lipoprotein metabolism: it's not just about sex hormones. *The Journal of clinical endocrinology and metabolism* 2011;96: 885-93.
8. del Rincón I, Battafarano DF, Restrepo JF, Erikson JM, Escalante A. Glucocorticoid dose thresholds associated with all-cause and cardiovascular mortality in rheumatoid arthritis. *Arthritis & rheumatology (Hoboken, NJ)* 2014;66: 264-72.
9. McGrath CM, Young SP. Lipid and Metabolic Changes in Rheumatoid Arthritis. *Curr Rheumatol Rep* 2015;17 9: 57.
10. Cuppen BV, Fu J, van Wietmarschen HA, Harms AC, Koval S, Marijnissen AC et al. Exploring the Inflammatory Metabolomic Profile to Predict Response to TNF-alpha Inhibitors in Rheumatoid Arthritis. *PLoS One* 2016;11 9: e0163087.
11. Hu C, van Dommelen J, van der Heijden R, Spijksma G, Reijmers TH, Wang M et al. RPLC-Ion-Trap-FTMS Method for Lipid Profiling of Plasma: Method Validation and Application to p53 Mutant Mouse Model. *Journal of Proteome Research* 2008;7: 4982-91.
12. Nurmohamed MT, Heslinga M, Kitas GD. Cardiovascular comorbidity in rheumatic diseases. *Nature Reviews Rheumatology* 2015;11: 693-704.
13. Hage MP, Al-Badri MR, Azar ST. A favorable effect of hydroxychloroquine on glucose and lipid metabolism beyond its anti-inflammatory role. *Therapeutic Advances in Endocrinology and Metabolism* 2014;5: 77-85.
14. Ishii I, Fukushima N, Ye X, Chun J. Lysophospholipid receptors: signaling and biology. *Annual review of biochemistry* 2004;73: 321-54.
15. Pruzanski W, Lambeau G, Lazdunski M, Cho W, Kopilov J, Kuksis A. Hydrolysis of minor glycerophospholipids of plasma lipoproteins by human group IIA, V and X secretory phospholipases A2. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* 2007;1771: 5-19.
16. Kougias P, Chai H, Lin PH, Lumsden AB, Yao Q, Chen C. Lysophosphatidylcholine and secretory phospholipase A2 in vascular disease: mediators of endothelial dysfunction and atherosclerosis. *Medical science monitor : international medical journal of experimental and clinical research* 2006;12: RA5-16.
17. Chisolm GM, 3rd, Chai Y. Regulation of cell growth by oxidized LDL. *Free Radic Biol Med* 2000;28 12: 1697-707.
18. Marathe GK, Silva AR, de Castro Faria Neto HC, Tjoelker LW, Prescott SM, Zimmerman Ga et al. Lysophosphatidylcholine and lyso-PAF display PAF-like activity derived from contaminating phospholipids. *Journal of lipid research* 2001;42: 1430-7.
19. Maiolino G, Rossitto G, Caielli P, Bisogni V, Rossi GP, Calò LA. The role of oxidized low-density lipoproteins in atherosclerosis: the myths and the facts. *Mediators of inflammation* 2013;2013: 714653.

## Supplementary Method

### Lipidomics profiling

For the detection of polar lipids, each 20  $\mu\text{L}$  serum was spiked with internal standard (ISTD) mix and lipids were extracted by 440  $\mu\text{L}$  methanol. After centrifugation, the supernatant was transferred to an Eppendorf tube and the solvent was evaporated using a Speedvac. The dried lipids were reconstituted in 145  $\mu\text{L}$  isopropanol, 0.1% formic acid and transferred to autosampler vials. In total 8  $\mu\text{L}$  was injected into ultra-performance liquid chromatography coupled to electrospray ionization-quadrupole time-of-flight (Agilent 6530 San Jose, USA) with an ACQUITY UPLC™ HSS T3 column (1.8  $\mu\text{m}$ , 2.1 $\times$ 100mm). Lipids were detected in full scan in the negative mode.

**Supplementary Table 1:** Baseline characteristics of patients in the cohort (n=281), split for GC non-users and users, females and males. P-values for comparisons among four subgroups were calculated by one-way ANOVA <sup>a</sup>, Chi-square tests <sup>b</sup>, and Kruskal wallis H tests <sup>c</sup>) based on distribution of the clinical parameter. Bold values indicate significant p-values (p < 0.05).

Item	Females		Males		p-value
	GC non-users (n = 136)	GC users (n = 77)	GC non-users (n = 26)	GC users (n = 42)	
Age, years, mean ± SD <sup>a</sup>	53.2 ± 12.8	53.8 ± 14.1	54.2 ± 12.2	57.0 ± 11.7	0.428
Body mass index, kg/m <sup>2</sup> , mean ± SD <sup>a</sup>	27.2 ± 5.8	27.0 ± 5.2	26.7 ± 3.4	25.0 ± 3.2	0.115
Current smoker, n (%) <sup>b</sup>	28 (20.6)	11 (14.3)	12 (46.2)	19 (45.2)	<b>&lt;0.001</b>
Positive RF, n (%) <sup>b</sup>	91 (66.9)	38 (49.4)	15 (57.7)	34 (81.0)	<b>0.004</b>
Positive ACPA, n (%) <sup>b</sup>	90 (66.2)	49 (63.6)	17 (65.4)	31 (73.8)	0.726
CRP, mg/dL, median (IQR) <sup>c</sup>	6.0 (2.7-15.0)	5.0 (3.0-16.5)	5.5 (2.3-12.0)	10 (2.3-35.3)	0.487
Use of other DMARDs, n (%) <sup>b</sup>					
Methotrexate, n (%) <sup>b</sup>	99 (72.8)	48 (62.3)	22 (84.6)	32 (76.2)	0.115
Leflunomide, n (%) <sup>b</sup>	17 (12.5)	9 (11.7)	2 (7.7)	7 (16.7)	0.737
Hydroxychloroquine, n (%) <sup>b</sup>	21 (15.4)	22 (28.6)	9 (34.6)	10 (23.8)	0.050
Use of non-DMARDs					
NSAIDs, n (%) <sup>b</sup>	94 (69.1)	39 (50.6)	20 (76.9)	22 (52.4)	<b>0.010</b>
Anticonception, n (%) <sup>b</sup>	10 (7.4)	5 (6.5)	0 (0.0)	0 (0.0)	0.162
Antidiabetics, n (%) <sup>b</sup>	5 (3.7)	4 (5.2)	1 (3.8)	1 (2.4)	0.892
Antihyperlipidemia, n (%) <sup>b</sup>	18 (13.2)	13 (16.9)	6 (23.1)	1 (2.4)	0.065
Bisphosphonates, n (%) <sup>b</sup>	34 (25.0)	58 (75.3)	3 (11.5)	30 (71.4)	<b>&lt;0.001</b>
DAS28, mean ± SD <sup>a</sup>	4.6 ± 1.2	4.6 ± 1.2	3.9 ± 1.1	4.4 ± 1.4	0.071
TJC, median (IQR) <sup>c</sup>	8.0 (2.3-12.0)	7.0 (2.0-15.0)	4.0 (1.0-6.3)	6.0 (0.0-14.0)	0.069
SJC, median (IQR) <sup>c</sup>	2.0 (0.0-4.0)	2.0 (0.0-4.0)	1.0 (0.0-3.0)	1.0 (0.0-4.3)	0.885
VAS-GH, mean ± SD <sup>a</sup>	58.6 ± 24.2	56.4 ± 22.7	49.2 ± 25.1	57.1 ± 23.3	0.324
ESR, mm/h, median (IQR) <sup>c</sup>	19.0 (11.0-37.0)	22.0 (6.5-39.0)	15.5 (6.0-34.3)	19.0 (9.5-47.5)	0.786

ACPA, anti citrullinated protein antibody, RF, c-reactive protein, DAS28, disease activity score based on a 28-joint count, DMARDs, disease-modifying antirheumatic drugs, ESR, erythrocyte sedimentation rate, IQR, interquartile range, NSAID, non-steroid anti-inflammatory drug, RF, rheumatoid factor, SD, standard deviation, SJC, swollen joint count, TJC, tender joint count, VAS-GH, visual analogue scale-general health.

**Supplementary Table 2:** List of detected metabolites in lipidomics profiling. The "sn1-" or "sn2-" prefix is used to indicate the position of the fatty acid chain esterified to the glycerol backbone of a lysophosphatidylcholine.

Lipid Class	Lipid Maps	Metabolite species	Amount (n)
<b>Free fatty acids (FA)</b>	FA01	C14:0, C14:1, C15:0, C16:0, C16:1, C17:0, C17:1, C18:0, C18:1, C18:2, C18:3- $\omega$ 3/ $\omega$ 6, C20:0, C20:1, C20:2, C20:3- $\omega$ 3/ $\omega$ 6, C20:3- $\omega$ 9, C20:4- $\omega$ 6, C20:5- $\omega$ 3, C22:4, C22:5- $\omega$ 3, C22:5- $\omega$ 6, C22:6, C24:0, C24:1	24
<b>Lysophosphatidylcholine (LPC)</b>	GP0105	<i>sn1</i> : C14:0, C15:0, C16:0, C16:1, C18:0, C18:1, C18:2, C18:3- $\omega$ 3/ $\omega$ 6, C19:0, C20:1, C20:2, C20:3- $\omega$ 3/ $\omega$ 6, C20:3- $\omega$ 9, C20:4, C20:5, C22:4, C22:5- $\omega$ 3, C22:5- $\omega$ 6, C22:6 <i>sn2</i> : C14:0, C16:0, C16:1, C18:0, C18:1, C18:2, C18:3- $\omega$ 3/ $\omega$ 6, C20:3- $\omega$ 3/ $\omega$ 6, C20:4, C20:5, C22:6	30
<b>Lysophosphatidylethanolamine (LPE)</b>	GP0205	C16:0, C18:0, C18:1, C18:2, C20:3- $\omega$ 3/ $\omega$ 6, C20:4, C20:5, C22:5- $\omega$ 3, C22:5- $\omega$ 6, C22:6	10
<b>Plasmalogen Lysophosphatidylcholine (LPC-O)<sup>a</sup></b>	GP0106	C16:0; C18:0; C18:1; C18:2	4

<sup>a</sup> The 'LPC-O' is used to identify a plasmalogen lysophosphatidylcholine, where the fatty acid chain is attached via a vinyl ether linkage to the glycerol backbone. The ' $\omega$ 3', ' $\omega$ 6', ' $\omega$ 9' are used to indicate double bond (C=C) at the third, sixth or ninth carbon atom from the end of the fatty acid chain.

**Supplementary Table 3:** Differences in 68 serum lipid levels between GC users and non-users in males and females, respectively. The mean differences between GC non-users and users are shown. P-values were calculated by means of an independent sample t-test.

Metabolite name	Males (GC user vs. GC non-user)		Females (GC user vs. GC non-user)	
	Mean <sub>non-user</sub> -Mean <sub>user</sub>	p value	Mean <sub>non-user</sub> -Mean <sub>user</sub>	p value
<i>sn1</i> -LPC (14:0)	-0.106	0.602	-0.444	0.001
<i>sn1</i> -LPC (15:0)	-0.116	0.633	-0.346	0.010
<i>sn1</i> -LPC (16:0)	0.027	0.796	-0.301	<0.001
<i>sn1</i> -LPC (16:1)	-0.181	0.422	-0.507	<0.001
<i>sn1</i> -LPC (18:0)	0.289	0.122	-0.148	0.304
<i>sn1</i> -LPC (18:1)	-0.091	0.512	-0.368	<0.001
<i>sn1</i> -LPC (18:2)	-0.283	0.027	-0.332	<0.001
<i>sn1</i> -LPC (18:3- $\omega$ 3/ $\omega$ 6)	-0.343	0.077	-0.393	0.004
<i>sn1</i> -LPC (19:0)	0.252	0.153	0.091	0.494
<i>sn1</i> -LPC (20:1)	0.120	0.568	-0.207	0.099
<i>sn1</i> -LPC (20:2)	-0.153	0.408	-0.395	0.004
<i>sn1</i> -LPC (20:3- $\omega$ 3/ $\omega$ 6)	0.152	0.342	-0.226	0.054
<i>sn1</i> -LPC (20:3- $\omega$ 9)	0.284	0.276	-0.198	0.144
<i>sn1</i> -LPC (20:4)	0.183	0.244	-0.2	0.032
<i>sn1</i> -LPC (20:5)	-0.417	0.062	-0.237	0.074
<i>sn1</i> -LPC (22:4)	0.228	0.359	-0.283	0.047
<i>sn1</i> -LPC (22:5- $\omega$ 3)	-0.161	0.369	-0.373	0.007
<i>sn1</i> -LPC (22:5- $\omega$ 6)	0.357	0.128	-0.099	0.455
<i>sn1</i> -LPC (22:6)	0.107	0.610	-0.021	0.882
<i>sn2</i> -LPC (14:0)	-0.157	0.405	-0.427	0.001
<i>sn2</i> -LPC (16:0)	0.178	0.341	-0.501	<0.001
<i>sn2</i> -LPC (16:1)	-0.237	0.308	-0.487	<0.001
<i>sn2</i> -LPC (18:0)	0.211	0.098	-0.115	0.231
<i>sn2</i> -LPC (18:1)	-0.114	0.559	-0.443	<0.001
<i>sn2</i> -LPC (18:2)	-0.395	0.052	-0.533	<0.001
<i>sn2</i> -LPC (18:3- $\omega$ 3/ $\omega$ 6)	-0.362	0.161	-0.411	0.001
<i>sn2</i> -LPC (20:3- $\omega$ 3/ $\omega$ 6)	0.226	0.281	-0.263	0.062
<i>sn2</i> -LPC (20:4)	0.215	0.384	-0.331	0.021
<i>sn2</i> -LPC (20:5)	-0.421	0.075	-0.278	0.046
<i>sn2</i> -LPC (22:6)	0.092	0.666	-0.07	0.616
LPC (O-16:0)	0.157	0.466	-0.343	0.020
LPC (O-18:0)	-0.019	0.928	-0.419	0.003
LPC (O-18:1)	0.190	0.410	-0.152	0.299
LPC (O-18:2)	0.080	0.707	-0.398	0.004
LPE (16:0)	-0.279	0.143	-0.505	0.001
LPE (18:0)	-0.102	0.588	-0.587	<0.001
LPE (18:1)	0.031	0.883	-0.449	0.001
LPE (18:2)	-0.349	0.093	-0.497	<0.001
LPE (20:3- $\omega$ 3/ $\omega$ 6)	-0.051	0.789	-0.469	0.001
LPE (20:4)	-0.062	0.785	-0.685	<0.001
LPE (20:5)	-0.463	0.029	-0.352	0.012
LPE (22:5- $\omega$ 3)	-0.328	0.089	-0.694	<0.001
LPE (22:5- $\omega$ 6)	0.294	0.215	-0.356	0.009
LPE (22:6)	-0.242	0.218	-0.512	0.001
FA (14:0)	0.089	0.654	-0.014	0.924

FA (14:1)	0.038	0.865	0.159	0.275
FA (15:0)	0.129	0.517	-0.027	0.855
FA (16:0)	-0.083	0.680	-0.006	0.966
FA (16:1)	0.113	0.600	0.042	0.770
FA (17:0)	-0.097	0.652	-0.057	0.692
FA (17:1)	0.041	0.852	0.023	0.872
FA (18:0)	-0.072	0.725	0.109	0.456
FA (18:1)	0.134	0.551	0.075	0.609
FA (18:2)	0.089	0.689	0.035	0.805
FA (18:3- $\omega$ 3/ $\omega$ 6)	0.190	0.364	0.009	0.950
FA (20:0)	0.039	0.878	-0.071	0.620
FA (20:1)	0.054	0.803	-0.013	0.925
FA (20:2)	0.011	0.969	-0.042	0.761
FA (20:3- $\omega$ 3/ $\omega$ 6)	0.377	0.109	-0.107	0.445
FA (20:3- $\omega$ 9)	0.500	0.021	-0.074	0.595
FA (20:4- $\omega$ 6)	0.250	0.212	-0.013	0.905
FA (20:5- $\omega$ 3)	-0.105	0.628	-0.098	0.463
FA (22:4)	0.237	0.234	-0.077	0.543
FA (22:5- $\omega$ 3)	-0.032	0.866	-0.143	0.271
FA (22:5- $\omega$ 6)	0.247	0.261	0.017	0.901
FA (22:6)	0.371	0.107	0.072	0.585
FA (24:0)	0.071	0.783	0.028	0.838
FA (24:1)	-0.013	0.951	0.016	0.901

**Supplementary Table 4:** Loading scores of 68 lipids in principle components analysis. Component loading scores of all lipids were presented for the first five components (with eigenvalue >1). Since the LPCs and LPEs showed high component loadings in one component (the first) and only with positive values, they all highly correlate in the same direction. Considering all LPCs and LPEs significantly different between GC users and non-users, all loading scores were  $\geq 0.438$ .

Metabolite name	Component				
	1	2	3	4	5
<i>sn1</i> -LPC (14:0)*	0.701	-0.253	0.057	0.147	-0.357
<i>sn1</i> -LPC (15:0)*	0.610	-0.212	-0.159	0.053	-0.066
<i>sn1</i> -LPC (16:0)*	0.880	-0.187	-0.077	-0.080	-0.009
<i>sn1</i> -LPC (16:1)*	0.773	-0.073	0.152	0.044	-0.319
<i>sn1</i> -LPC (18:0)*	0.744	-0.172	-0.249	-0.082	0.295
<i>sn1</i> -LPC (18:1)*	0.854	-0.228	0.023	0.131	0.135
<i>sn1</i> -LPC (18:2)*	0.701	-0.297	-0.036	0.181	0.429
<i>sn1</i> -LPC (18:3- $\omega$ 3/ $\omega$ 6)*	0.694	-0.295	0.047	0.344	-0.100
<i>sn1</i> -LPC (19:0)	0.421	-0.235	-0.315	-0.029	0.291
<i>sn1</i> -LPC (20:1)	0.336	-0.041	-0.252	-0.132	0.203
<i>sn1</i> -LPC (20:2)*	0.800	-0.156	0.126	-0.108	0.108
<i>sn1</i> -LPC (20:3- $\omega$ 3/ $\omega$ 6)	0.665	-0.074	0.404	-0.312	-0.133
<i>sn1</i> -LPC (20:3- $\omega$ 9)	0.652	-0.120	0.345	-0.165	-0.213
<i>sn1</i> -LPC (20:4)*	0.674	0.085	0.154	-0.492	0.032
<i>sn1</i> -LPC (20:5)	0.543	0.016	-0.668	0.069	-0.296
<i>sn1</i> -LPC (22:4)*	0.676	-0.040	0.539	-0.326	0.026
<i>sn1</i> -LPC (22:5- $\omega$ 3)*	0.785	-0.040	-0.219	-0.022	0.014
<i>sn1</i> -LPC (22:5- $\omega$ 6)	0.556	0.003	0.569	-0.405	-0.068
<i>sn1</i> -LPC (22:6)	0.534	0.160	-0.574	-0.254	-0.084
<i>sn2</i> -LPC (14:0)*	0.674	-0.226	0.057	0.173	-0.343
<i>sn2</i> -LPC (16:0)*	0.832	-0.214	-0.050	-0.057	-0.051
<i>sn2</i> -LPC (16:1)*	0.713	-0.063	0.183	0.026	-0.314
<i>sn2</i> -LPC (18:0)	0.761	-0.164	-0.235	-0.079	0.283
<i>sn2</i> -LPC (18:1)*	0.789	-0.219	0.037	0.112	0.121
<i>sn2</i> -LPC (18:2)*	0.693	-0.298	-0.018	0.160	0.439
<i>sn2</i> -LPC (18:3- $\omega$ 3/ $\omega$ 6) *	0.614	-0.218	0.049	0.322	-0.104
<i>sn2</i> -LPC (20:3- $\omega$ 3/ $\omega$ 6)	0.670	-0.089	0.385	-0.302	-0.079
<i>sn2</i> -LPC (20:4)*	0.689	0.048	0.175	-0.463	0.024
<i>sn2</i> -LPC (20:5)*	0.548	0.031	-0.643	0.089	-0.315
<i>sn2</i> -LPC (22:6)	0.563	0.126	-0.551	-0.251	-0.070
LPC (O-16:0)*	0.632	-0.225	-0.219	-0.301	0.408
LPC (O-18:0)*	0.542	-0.258	-0.252	-0.242	0.411
LPC (O-18:1)*	0.536	-0.172	-0.241	-0.290	0.544
LPC (O-18:2)*	0.549	-0.083	-0.283	-0.209	0.406
LPE (16:0)*	0.681	-0.248	0.066	0.176	-0.278
LPE (18:0)*	0.746	0.039	0.082	0.184	-0.072
LPE (18:1)*	0.460	-0.306	0.210	0.558	0.173
LPE (18:2)*	0.513	-0.374	0.219	0.588	0.217
LPE (20:3- $\omega$ 3/ $\omega$ 6)*	0.601	-0.099	0.512	0.237	-0.167
LPE (20:4)*	0.667	-0.050	0.375	0.126	0.044
LPE (20:5)*	0.483	-0.075	-0.496	0.458	-0.328
LPE (22:5- $\omega$ 3)*	0.736	-0.073	0.161	0.340	-0.156

LPE (22:5- $\omega$ 6)*	0.438	0.015	0.672	-0.080	-0.177
LPE (22:6)*	0.604	0.179	-0.323	0.128	-0.235
FA (14:0)	0.192	0.883	0.066	0.221	-0.006
FA (14:1)	0.141	0.841	0.023	0.089	-0.061
FA (15:0)	0.215	0.733	0.102	0.210	0.065
FA (16:0)	0.206	0.770	0.190	0.294	0.183
FA (16:1)	0.164	0.921	0.067	0.020	0.000
FA (17:0)	0.221	0.744	0.069	0.239	0.151
FA (17:1)	0.219	0.910	0.031	0.068	0.043
FA (18:0)	0.120	0.378	0.218	0.392	0.259
FA (18:1)	0.106	0.941	0.057	0.073	0.130
FA (18:2)	0.137	0.894	0.012	0.087	0.206
FA (18:3- $\omega$ 3/ $\omega$ 6)	0.181	0.853	-0.044	0.183	0.119
FA (20:0)	0.151	0.154	0.105	0.336	0.209
FA (20:1)	0.115	0.921	-0.029	0.071	0.160
FA (20:2)	0.155	0.781	0.008	-0.129	0.112
FA (20:3- $\omega$ 3/ $\omega$ 6)	0.315	0.732	0.216	-0.184	-0.111
FA (20:3- $\omega$ 9)*	0.285	0.577	0.152	-0.208	-0.124
FA (20:4- $\omega$ 6)	0.256	0.615	-0.019	-0.349	0.074
FA (20:5- $\omega$ 3)	0.339	0.470	-0.686	0.085	-0.244
FA (22:4)	0.199	0.861	0.266	-0.164	-0.002
FA (22:5- $\omega$ 3)	0.312	0.846	-0.275	0.076	-0.065
FA (22:5- $\omega$ 6)	0.248	0.783	0.271	-0.237	-0.142
FA (22:6)	0.242	0.629	-0.494	-0.076	-0.163
FA (24:0)	-0.190	-0.080	0.319	0.392	0.553
FA (24:1)	0.228	0.637	-0.163	0.125	0.118

\* Metabolites were also significantly different in independent t-tests between GC users and non-users, in either females or males (Figure 1, Supplementary Table 3).

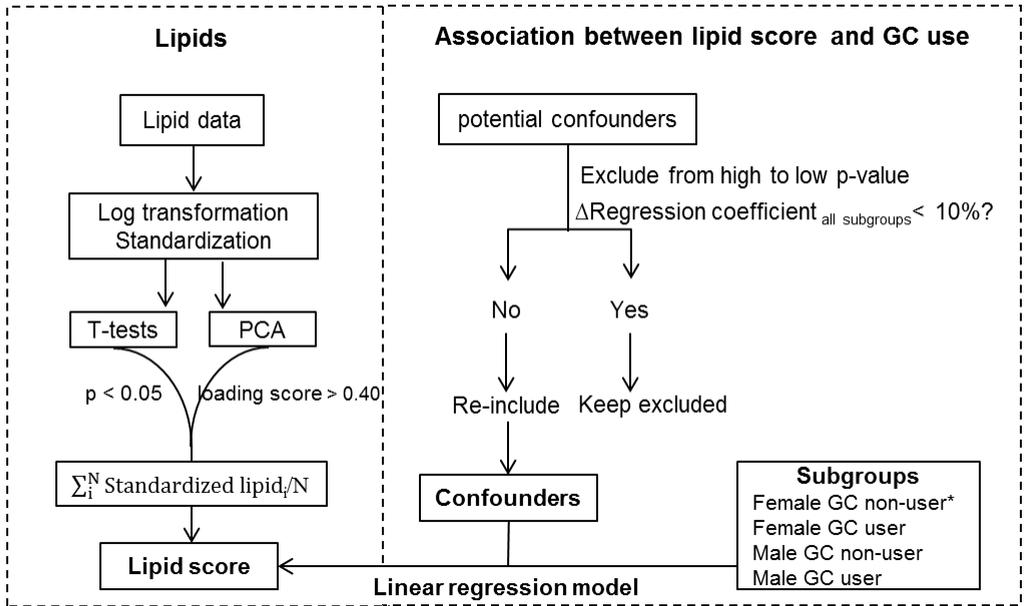
**Supplementary Table 5:** The final regression model investigating the association between gender and GC use on the lysophospholipid score.

<b>Variables</b>	<b>Coefficients<sup>a</sup></b>	<b>95%-confidence interval</b>	<b>p value</b>
<b>Male patients no GC use (reference group)</b>	0.149	(-0.491 to 0.789)	0.649
<b>Male patients GC use</b>	0.122	(-0.160 to 0.405)	0.397
<b>Female patients no GC use</b>	-0.310	(-0.604 to -0.015)	0.041
<b>Female patients GC use</b>	0.089	(-0.212 to 0.390)	0.565
<b>Age</b>	0.001	(-0.008 to 0.007)	0.937
<b>Menopausal status</b>	0.286	(0.053 to 0.519)	0.017
<b>BMI (kg/m<sup>2</sup>)</b>	0.006	(-0.009 to 0.020)	0.447
<b>RF positivity</b>	-0.247	(-0.432 to -0.063)	0.009
<b>ACPA positivity</b>	0.182	(-0.008 to 0.372)	0.061
<b>CRP (log transformed, mg/dL)</b>	-0.122	(-0.192 to -0.052)	0.001
<b>ESR (log transformed, mm/h)</b>	0.058	(-0.031 to 0.147)	0.202
<b>DAS28</b>	-0.014	(-0.078 to 0.050)	0.663
<b>Concomitant methotrexate</b>	-0.023	(-0.181 to 0.134)	0.773
<b>Concomitant hydroxychloroquin</b>	-0.180	(-0.347 to -0.013)	0.035
<b>Antidiabetics</b>	-0.132	(-0.493 to 0.229)	0.474
<b>NSAIDs</b>	-0.068	(-0.215 to 0.078)	0.362

<sup>a</sup> Coefficients indicate the changes in mean lysophospholipid score after correcting for confounders—BMI, menopausal status, RF positive, log-transformed CRP, log transformed ESR, DAS28, and concomitant drugs (methotrexate, hydroxychloroquine, anti-diabetic drug and NSAIDs). BMI, body mass index; CRP, C-reactive protein; GC, glucocorticoid; RF, rheumatoid factor; ACPA, anti citrullinated protein antibody; ESR, erythrocyte sedimentation rate; DAS28, disease activity score based on a 28-joint count; NSAIDs, non-steroid anti-inflammatory drugs

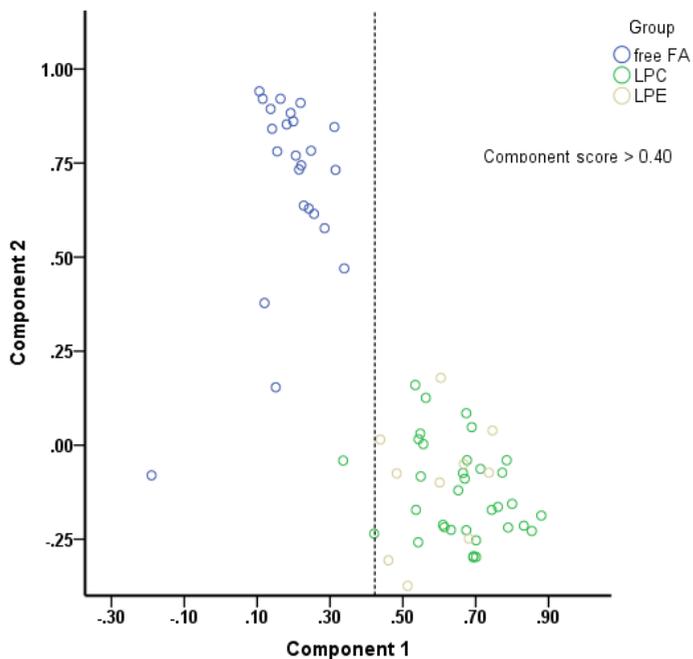
**Supplementary Figure 1:** Schematic overview of the statistical analyses.

the subgroup "Female-GC non-user" is set as the reference group in the linear regression analysis.



**Supplementary Figure 2:** Graphical representation of loading scores of 68 lipids in component 1 and 2 by PCA analysis. All the lysophospholipids (green and yellow circles) are with loading scores > 0.40 in component 1 (on the right side of the dash line).

FA, fatty acid; LPC, lysophosphatidylcholines; LPE, lysophosphatidylethanolamine





## **Chapter 10**

### **Polymorphisms in the multidrug-resistance 1 gene related to glucocorticoid response in rheumatoid arthritis treatment**

Bart V.J. Cuppen, Katerina Pardali, Maarten C. Kraan, Anne C.A. Marijnissen, Linda Yrlid, Marita Olsson, Johannes W.J. Bijlsma, Floris P.J.G. Lafeber, Ruth D.E. Fritsch-Stork

## **ABSTRACT**

### **Objective**

A substantial proportion of rheumatoid arthritis (RA)-patients experience an insufficient response to glucocorticoids, an important therapeutic agent in RA. The multidrug resistance 1 (MDR1) gene product P-glycoprotein (P-gp) is an efflux pump that actively transports substrates such as glucocorticoids, out of the cell. We investigated if the variation in response might be explained by single nucleotide polymorphisms (SNPs) in the MDR1 gene.

### **Methods**

RA-patients treated with intravenous methylprednisolone pulses (n=18) or oral prednisone/prednisolone (n=22) were included in a prospective cohort and clinical response was measured after 5 and 30 days, respectively. The C1236T, G2677A/T, and C3435T SNPs were determined and the functionality of P-gp was assessed by flow cytometry (Rhodamine efflux assay).

### **Results**

Carriage of the G2677AT SNP was significantly associated with response (OR=6.18, p=0.035), the other SNPs showed trends. Stratified for received treatment, the effect was only present in methylprednisolone treated patients. Mutant allele carriage significantly decreased functionality of P-gp in B-cells, though had a smaller impact in other PBMC subtypes.

### **Conclusion**

Carriage of a MDR1 SNP was related to a response to methylprednisolone in this study, which suggests that RA-patients carrying wild type alleles might benefit from P-gp inhibition or administration of glucocorticoid analogues that are non-P-gp substrates.

## **INTRODUCCION**

Rheumatoid arthritis (RA) is a chronic, disabling disease that mainly affects the synovial joints. Glucocorticoids (GC) compose a class of drugs that have an important place in RA treatment [1], however, a substantial proportion of patients experiences insufficient response.

The multidrug resistance 1 (MDR1) gene product, the P-glycoprotein (P-gp), is an efflux pump that actively transports substrates such as drugs, out of the cell. The ability to regulate intracellular substrate concentration, depends on both the expression and the functionality (i.e. recognition of substrates and transport effectiveness) of P-gp [2]. In treatment refractory RA, P-gp overexpression in lymphocytes is believed to play a substantial role, as it has been shown that overexpression reduces intracellular concentrations of immunosuppressants that are P-gp substrates [3]. Single nucleotide polymorphisms (SNPs) have been linked to P-gp activity and are suspected to influence inter-individual variation in response to certain treatments [4, 5]. Because GCs are known P-gp substrates [6, 7], differences in GC-response might be explained by carriage of polymorphisms in the MDR1 gene. We therefore explored the possibility that carriers of mutant alleles in the MDR1 gene efflux GCs less effectively, and subsequently experience a better clinical response.

## **METHODS**

### **Patients and sample collection**

RA-patients treated with GCs at our department of Rheumatology were included in a prospective cohort. Patients eligible for intravenous (IV) GC pulse treatment according to their treating rheumatologist, were given three doses of 1000mg IV methylprednisolone on alternate days, as this approach has been proven to be quick, effective (in short and long term) and safe [8-10]. RA-patients with an indication for oral GC treatment were treated with prednisone or prednisolone doses between 7.5-10mg for at least one month. Based on our clinical experience and previous research on IV pulse treatment [8], the optimal time point to measure (maximum) treatment effect was considered five days after start. Treatment response in orally treated patients was assessed at day 30 which was

considered sufficient as a response as early as two weeks is already a strong indicator of the long term clinical outcome [11]. DAS28 response was evaluated according to the European league against Rheumatism (EULAR) response criteria [12]. EULAR good response, i.e. a DAS  $\geq 1.2$  and DAS28  $< 3.2$ , was compared to the combination of moderate and non-response considering the intensity of the treatment, especially in the IV arm.

Whole blood was collected in Li-Heparin tubes. Peripheral blood mononuclear cells (PBMCs) were isolated and stored in liquid nitrogen until use. The study was approved by the medical ethics committee of the University Medical Center Utrecht and was conducted in accordance with the Helsinki Declaration. All patients signed informed consent.

### **Genotyping of MDR1**

Genomic DNA was extracted from  $10^5$  PBMCs using QIAamp DNA blood mini kit (Cat. No 51104, Qiagen, Hilden, Germany) according to manufacturer's instructions. C1236T (rs1128503), G2677A/T (rs2032582), and C3435T (rs1045642) SNPs were genotyped using TaqMan allelic discrimination assay technology (Applied Biosystems, Foster City, California, USA) on a QuantStudio 7 Flex Real-Time PCR System (Thermo Fisher Scientific Inc, Foster City, California, USA), according to manufacturer's instructions. The analyses were run in duplicate. The mutant allele frequency of the study cohort (observed frequency) was compared with the Central European occurrence (expected allele frequency, derived from 1000 Genomes catalog [13]).

### **Functionality of P-gp**

The functionality of P-gp was assessed by flow cytometry, using a Rhodamine efflux assay.  $4.0 \times 10^5$  cryo-preserved PBMCs were incubated with Rhodamine 123 (Sigma-Aldrich, St. Louis, USA) at a final concentration of 1  $\mu\text{g/ml}$  for 30 minutes at 4 °C and washed twice. After dividing cells in two wells, they were incubated in the presence (control) or absence of the MDR1-inhibitor Elacridar (Astrazeneca, Mölndal, Sweden). The cells were stained with antibodies to determine subsets of PBMCs: CD3, CD4, CD8, CD14, CD16, CD19, CD196 and CD56 (all BD Bioscience, San Jose, USA). Flow cytometry was performed on a 4 laser LSR Fortessa (BD Bioscience, San Jose, USA) and samples were analyzed using FlowJo (Tree

Star Inc. Ashland, USA). Only baseline samples (i.e. before start of therapy) with a clear FSC vs SSC profile containing living PBMCs were used for data analysis. As additive test for viability and proper labeling with Rhodamine-123, all samples were excluded that had >5% low Rhodamine-123 containing cells in the elacridar treated sample. After gating of living cells in the FSC vs SSC plot, duplicates and cell-aggregates were removed using FSC-A vs FSC-H plot. The frequency of cells effluxing Rhodamine-123 in the uninhibited samples was used as a value for transport functionality.

### **Statistical analysis**

Associations between each SNP and patient's response was analyzed using logistic regression models, in which each investigated SNP was tested while adjusted for treatment received (IV/oral). SNPs were first analyzed in a dominant model in which the SNP was coded dichotomously (wildtype/mutant allele carrier) and subsequently in an additive model (linear covariate coded: zero mutant alleles (wildtype)/ one mutant allele (heterozygous)/ two mutant alleles (homozygous mutant)). All analyses were performed in SPSS 21 (IBM Corp., Armonk, NY).

### **RESULTS**

Baseline characteristics of patients treated with IV methylprednisolone (n=18) and oral prednisone/prednisolone (n=22) are shown in Table 1. There were no differences between the cohorts, apart from a higher DAS28 at baseline in line with the difference in clinical indication for the two treatment regimens.

**Table 1:** Baseline characteristics for intravenous and oral GC treated cohorts. P-values were calculated by Fisher exact or t-test for resp. dichotomous and continuous variables.

Item	IV cohort (n=18)	Oral cohort (n=22)	P-value
Female gender, n (%)	13 (72.2)	18 (81.8)	0.71
Age, mean ( $\pm$ sd)	61.4 ( $\pm$ 17.7)	59.8 ( $\pm$ 15.1)	0.76
DMARD naïve, n (%)	3 (16.7)	6 (27.3)	0.70
Biological use, n (%)	2 (11.1)	4 (18.2)	0.68
MTX use, n (%)	9 (50.0)	12 (54.5)	1.00
LEF use, n (%)	2 (11.1)	1 (4.5)	0.58
AZA use, n (%)	0 (0.0)	1 (4.5)	1.00
SSZ use, n (%)	1 (5.6)	1 (4.5)	1.00
HQC use, n (%)	2 (11.1)	3 (13.6)	1.00
DAS28 Baseline, mean ( $\pm$ sd)	6.3 ( $\pm$ 1.1)	5.2 ( $\pm$ 1.5)	0.01

AZA, azathioprine, DAS28, disease activity score based on 28-joint count, DMARD, disease modifying anti-rheumatic drugs, HCQ, hydroxychloroquine, LEF, leflunomide, MTX, methotrexate, SSZ, sulfasalazine.

Since concomitantly used disease-modifying antirheumatic drugs (DMARDs) might confound the investigated relationship of GCs and response, we additionally investigated the distribution of concomitant DMARDs use among the responders and non-responders (Supplementary Table 1). However, no effect of any DMARD on (non-)response was seen.

### Mutant allele frequencies

The observed allele frequency was lower than expected based on the general population, however, these differences did not reach statistical significance: 0.36 for T-allele of C1236T (0.43 expected, binominal test of proportion  $p=0.10$ ), 0.37 for A or T-allele of G2677AT (0.44 expected,  $p=0.11$ ) and 0.50 for T-allele C3435T (0.57 expected,  $p=0.11$ ).

### Genotype and Response to GC therapy

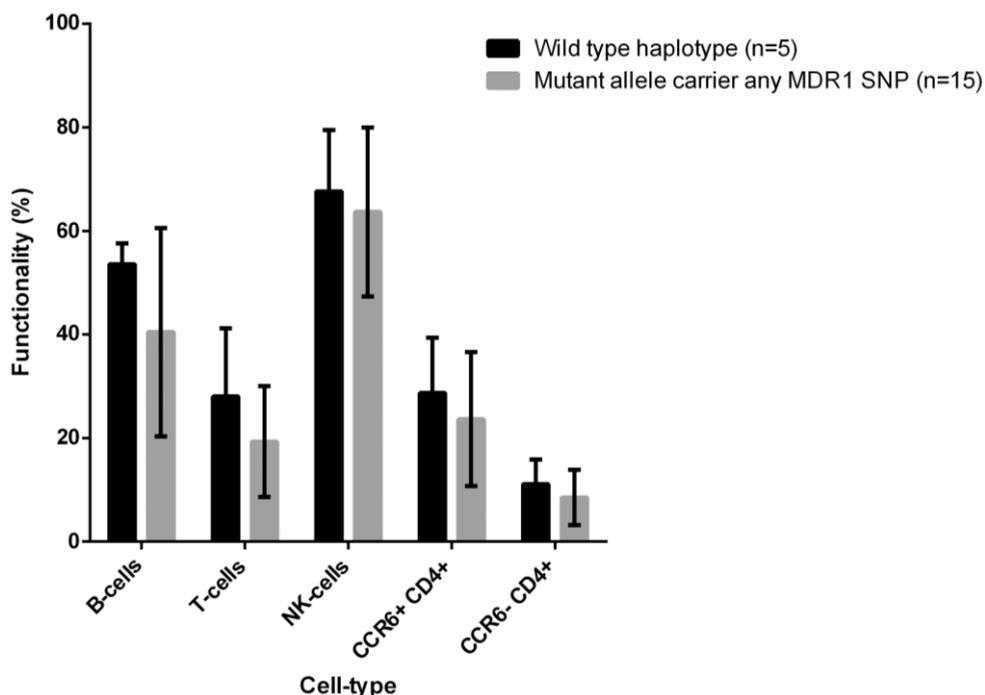
The dominant model revealed an OR of 6.18 for G2677AT polymorphism carriers and a trend for the other SNPs (Table 2). In the additive model a trend for all SNPs and mutant allele carriers was found (data not shown). Stratified for the received treatment, wild type carriers for all SNP in the IV cohort failed to achieve response, whereas 43-50% of the mutant allele carriers achieved a response ( $p=0.05$  for G2677AT). For oral treatment no differences in SNPs and response rates were observed.

**Table 2:** SNPs and clinical response to IV and oral GC treatment. Wild-type (for each respective allele) versus mutant allele carriers were compared in a logistic regression analysis corrected for treatment received. Per SNP the odds ratio (OR) and related p-value are presented. In the subsequent columns, the absolute number of responders and non-responders per treatment is shown for wildtype and mutant allele carriers, including the overall p-value (Fisher exact test). For the G2677AT polymorphism there was a significant association between mutant allele carriage and response, though this association was restricted to the IV treated patients. Trends for the other SNPs and response to IV treatment were seen.

SNP	Overall OR	P-value for OR	Treatment	Genotype	Non-resp n (%)	Resp n (%)	p-value
<b>C1236T</b>	1.96	0.35	IV	CC (n=5)	5 (100)	0 (0)	0.11
				CT/TT (n=13)	7 (54)	6 (46)	
			Oral	CC (n=9)	5 (56)	4 (44)	0.66
				CT/TT (n=13)	8 (62)	5 (39)	
<b>G2677AT</b>	6.18	<b>0.04</b>	IV	GG (n=6)	6 (100)	0 (0)	<b>0.05</b>
				GT/TT (n=12)	6 (50)	6 (50)	
			Oral	GG (n=8)	6 (75)	2 (25)	0.38
				GT/GA/AT/TT (n=14)	7 (50)	7 (50)	
<b>C3435T</b>	1.69	0.51	IV	CC (n=3)	3 (100)	0 (0)	0.52
				CT/TT (n=15)	9 (60)	6 (40)	
			Oral	CC (n=7)	4 (57)	3 (43)	1.00
				CT/TT (n=15)	9 (60)	6 (40)	

### Functionality of P-gp

Of the originally 40 samples, 18 had to be excluded for P-gp functionality testing due to viability <40% and an additional two as they were taken after therapy initiation. Mutant allele carriers of all three SNPs showed lower functionality in B-cells (p=0.03), and for other cell sets a small decrease in functionality was observed (all p>0.10, Fig 1).



**Fig 1:** Functionality of P-gp for different subsets of PBMCs in %. Shown are the mean values and the standard deviations between wild type carriers for all SNPs (haplotype carriers, n=5) and mutant allele carriers for any MDR1 SNP (n=15). CCR6+ CD4+ cells include most of the Th17 cells and subpopulations of regulatory- and memory T-cells expressing CCR6+. Mutant allele carriers showed a decreased functionality of B-cells (t-test, p=0.03) and minor decreased functionality for other cells subsets (all p>0.10).

## DISCUSSION

The present study investigated the role of polymorphisms in the MDR1 gene in the response to IV and oral GC therapy in RA, and showed that the response to IV methylprednisolone is significantly better in patients carrying the G2677AT polymorphism compared to the wild type carriers, with trends for the other SNPs. Mutant allele carriage for any MDR1 SNP was associated with a significant decreased P-gp functionality in B-cells, whereas the impact on other PBMC subtypes was smaller but still present. Non-response to methylprednisolone, might therefore be (partially) explained by the role of polymorphisms in the MDR1 gene and subsequent effect on P-gp activity

The observed difference in effect of polymorphisms on the GC response in IV methylprednisolone compared to prednisone/prednisolone treated patients, could be explained by two mechanisms. First, a profound difference in dependency on P-gp among GC analogues has been shown [14]. Methylprednisolone has the highest transport efficiency by P-gp of all GCs, and thus is transported more efficiently than prednisolone (i.e. active component of prednisone). Consequently, an impaired function of P-gp due to a mutant allele exerts more impact in methylprednisolone compared with prednisone/prednisolone treated patients, which is in line with our results. Second, different substrate concentrations might lead to different effects on (dys)function. Whereas at low dose therapy (oral cohort), substrate influx and P-gp excretion may be in equilibrium, a drastic increase in substrate concentration such as in IV pulse treatment could result in a relative dysfunctionality of P-gp. This hypothesis fits the observations from *in vitro* work demonstrating differences in functionality between wild type and polymorphism carriers at high substrate concentration (2.6  $\mu\text{M}$  [15] and 5.0  $\mu\text{M}$  [4]) but not at lower substrate concentrations (0.4  $\mu\text{M}$  [16]). In this study a substrate concentration of 2.6  $\mu\text{M}$  was used, which showed lower efflux in B-cells though smaller differences in other cell subsets.

Several other studies have focused on the relation of SNPs and response to GCs treatments in inflammatory diseases. Two studies investigated the C3435T polymorphism in RA and the effect on the combination therapy of methotrexate and GCs, which yielded better response rates for C3435T mutant allele carriers [17, 18]. Carriers of the G2677AT polymorphism in immune thrombocytopenia (ITP) have a better response to oral GCs, whereas the carriage of mutant alleles for C1236T and C3435T did not correlate with response [19]. In inflammatory bowel diseases G2677AT did not relate to response, but C1236T and C3435T did [20]. The differences in results among these studies can partially be explained by factors related to demographics, disease, treatment and study design (including statistical power of the different studies), and therefore do not exclude the possibility of a shared underlying pathway related to one or more SNPs.

Our study has its limitations with respect to the number of patients that were included, resulting in a reduced power for detecting effects and hampering correction for multiple testing. Therefore, the analyses should be considered exploratory, and until validated in a bigger cohort the results should be interpreted in that context. It should also be noted, that although GC therapy was the subject of our study and concomitant treatments were stable, there were - albeit not significant - differences in used concomitant DMARDs between the cohorts and between responders and non-responders. Because of the low sample size, however, it was not possible to correct for these differences in multivariable analyses. In this study, the possibility of (extra-cytosolic) non-genomic effects to GC treatment were not considered. However, such effects are expected to appear more rapidly than 5/30 days and are therefore not likely to have influenced the observations. The chosen time intervals to evaluate the clinical response might have influenced the magnitude of the observed response, yet, the time points were based on empirical observations and were considered clinically most relevant. Another limitation is the low viability of PBMCs after thawing, possibly related to factors such as transport or the age of the cells. The dead cells could have disturbed the results of the functionality assay, which would explain the limited differences in P-gp functionality between the polymorphisms.

## **CONCLUSION**

Our data suggest that RA-patients with wild type polymorphisms could potentially benefit from concomitant MDR1-inhibition (such as by cyclosporine, tacrolimus and hydroxychloroquine [3]), administration of GC analogues that are less sensitive to P-gp transport (or not a substrate for P-gp) or treatments that are not influenced by P-gp, e.g. extracellular acting treatments such as tumor necrosis factor alpha inhibitors. Additional work needs to be undertaken using larger patient cohorts to better clarify the association between MDR1 SNPs, P-gp activity and response to GCs.

## REFERENCES

1. Kirwan JR, Bijlsma JW, Boers M, Shea BJ. Effects of glucocorticoids on radiological progression in rheumatoid arthritis. *Cochrane Database Syst Rev* 2007 1: CD006356.
2. Hoffmeyer S, Burk O, von RO, Arnold HP, Brockmoller J, John A et al. Functional polymorphisms of the human multidrug-resistance gene: multiple sequence variations and correlation of one allele with P-glycoprotein expression and activity in vivo. *Proc Natl Acad Sci USA* 2000;97 7: 3473-8.
3. Tsujimura S, Tanaka Y. Disease control by regulation of P-glycoprotein on lymphocytes in patients with rheumatoid arthritis. *World J Exp Med* 2015;5 4: 225-31.
4. Salama NN, Yang Z, Bui T, Ho RJ. MDR1 haplotypes significantly minimize intracellular uptake and transcellular P-gp substrate transport in recombinant LLC-PK1 cells. *J Pharm Sci* 2006;95 10: 2293-308.
5. Tsujimura S, Saito K, Nawata M, Nakayamada S, Tanaka Y. Overcoming drug resistance induced by P-glycoprotein on lymphocytes in patients with refractory rheumatoid arthritis. *Ann Rheum Dis* 2008;67 3: 380-8.
6. Ueda K, Okamura N, Hirai M, Tanigawara Y, Saeki T, Kioka N et al. Human P-glycoprotein transports cortisol, aldosterone, and dexamethasone, but not progesterone. *J Biol Chem* 1992;267 34: 24248-52.
7. Webster JI, Carlstedt-Duke J. Involvement of multidrug resistance proteins (MDR) in the modulation of glucocorticoid response. *J Steroid Biochem Mol Biol* 2002;82 4-5: 277-88.
8. Jacobs JW, Geenen R, Evers AW, van Jaarsveld CH, Kraaijmaat FW, Bijlsma JW. Short term effects of corticosteroid pulse treatment on disease activity and the wellbeing of patients with active rheumatoid arthritis. *Ann Rheum Dis* 2001;60 1: 61-4.
9. Forster PJ, Grindulis KA, Neumann V, Hubball S, McConkey B. High-dose intravenous methylprednisolone in rheumatoid arthritis. *Ann Rheum Dis* 1982;41 5: 444-6.
10. de Jong PH, Hazes JM, Barendregt PJ, Huisman M, van Zeben D, van der Lubbe PA et al. Induction therapy with a combination of DMARDs is better than methotrexate monotherapy: first results of the tREACH trial. *Ann Rheum Dis* 2013;72 1: 72-8.
11. de Jong PH, Quax RA, Huisman M, Gerards AH, Feelders RA, de Sonnaville PB et al. Response to glucocorticoids at 2 weeks predicts the effectiveness of DMARD induction therapy at 3 months: post hoc analyses from the tREACH study. *Ann Rheum Dis* 2013;72 10: 1659-63.
12. van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;41 10: 1845-50.
13. Institute TEB. 1000 Genomes; A Deep Catalog of Human Genetic Variation. <http://www.1000genomes.org/> Accessed 23 July 2015.
14. Yates CR, Chang C, Kearbey JD, Yasuda K, Schuetz EG, Miller DD et al. Structural determinants of P-glycoprotein-mediated transport of glucocorticoids. *Pharm Res* 2003;20 11: 1794-803.
15. Hitzl M, Drescher S, van der Kuip H, Schaffeler E, Fischer J, Schwab M et al. The C3435T mutation in the human MDR1 gene is associated with altered efflux of the P-glycoprotein substrate rhodamine 123 from CD56+ natural killer cells. *Pharmacogenetics* 2001;11 4: 293-8.
16. Oselin K, Gerloff T, Mrozikiewicz PM, Pahkla R, Roots I. MDR1 polymorphisms G2677T in exon 21 and C3435T in exon 26 fail to affect rhodamine 123 efflux in peripheral blood lymphocytes. *Fundam Clin Pharmacol* 2003;17 4: 463-9.
17. Pawlik A, Wrzesniewska J, Fiedorowicz-Fabrycy I, Gawronska-Szklarz B. The MDR1 3435 polymorphism in patients with rheumatoid arthritis. *Int J Clin Pharmacol Ther* 2004;42 9: 496-503.
18. Drozdziak M, Rudas T, Pawlik A, Kurzawski M, Czerny B, Gornik W et al. The effect of 3435C>T MDR1 gene polymorphism on rheumatoid arthritis treatment with disease-modifying antirheumatic drugs. *Eur J Clin Pharmacol* 2006;62 11: 933-7.
19. Xuan M, Li H, Fu R, Yang Y, Zhang D, Zhang X et al. Association of ABCB1 gene polymorphisms and haplotypes with therapeutic efficacy of glucocorticoids in Chinese patients with immune thrombocytopenia. *Hum Immunol* 2014;75 4: 317-21.
20. Yang QF, Chen BL, Zhang QS, Zhu ZH, Hu B, He Y et al. Contribution of MDR1 gene polymorphisms on IBD predisposition and response to glucocorticoids in IBD in a Chinese population. *J Dig Dis* 2015;16 1: 22-30.

**Supplementary Table 1:** Concomitant treatments among responders and non-responders to GC therapy.

	<b>Non-response (n=25)</b>	<b>Good response (n=15)</b>	<b>p-value</b>
<b>Biological, n (%)</b>	4 (16.0)	2 (13.3)	1.00
<b>MTX use, n (%)</b>	13 (52.0)	8 (53.3)	1.00
<b>LEF use, n (%)</b>	3 (12.0)	0 (0.0)	0.28
<b>AZA use, n (%)</b>	1 (4.0)	0 (0.0)	1.00
<b>SSZ use, n (%)</b>	1 (4.0)	1 (6.7)	1.00
<b>HQC use, n (%)</b>	2 (8.0)	3 (20.0)	0.35

AZA, azathioprine, HCQ, hydroxychloroquine, LEF, leflunomide, MTX, methotrexate, SSZ, sulfasalazine





## **Chapter 11**

### **Summary of results and general discussion**

## SUMMARY OF RESULTS IN THIS THESIS

This thesis aimed to identify the most valuable (combination of) predictors of response to therapy, in particular to tumour necrosis factor-alpha inhibitors (TNFi's), in rheumatoid arthritis (RA), by addressing the following questions:

*What is known about predictive markers of response to treatment in RA?*

Many studies to date have been performed that aimed to identify (bio)markers predictive of response to biological disease-modifying antirheumatic drugs (bDMARDs) in RA, of which the results were compiled in **chapter 2**. After evaluation, none of the single markers proposed as predictors and studied in multiple cohorts (replicated) consistently predicted response with an acceptable predictive ability. Many predictors were considered to have a more than acceptable predictive ability, yet were never replicated in subsequent studies. Validation of potential biomarkers remains key in prognostic research and should preferably be demonstrated in a separate additional cohort.

*Who will (dis)continue TNFi treatment?*

**Chapter 3** investigated treatment success in BiOCURA, defined as continuation or discontinuation of TNFi treatment in the first year after start. We noticed many discontinuations due to inefficacy and side effects of treatment, which were associated with the number of previously failed bDMARDs and with current smoking, in line with existing literature. The number of previously used bDMARDs as a predictor is likely a result of the selection of refractory RA patients; whereas smoking and its increased discontinuation rate may be related to more severe disease, pharmacological effects of smoking or to socioeconomic status. However, since the predictive abilities found on group level are too low for reliable individual prediction, these results are not useful to further personalised treatment decisions, but rather provide insight in factors that directly or indirectly influence treatment success.

*Can we find better biomarkers:*

1) *By using microRNAs?*

In **Chapter 4** it was observed that serum microRNAs (miRNAs) were not able to predict response to therapy consistently, as were the miRNAs that were previously proposed in two (unvalidated) studies. Cancer studies teach us that false positive results in exploration of circulating miRNAs are lurking, and our additional analyses provided possible explanations for the failed validation in this and other studies predicting response. To increase the robustness of proposed miRNA biomarkers, future studies should take the influence of clinical parameters into account - in order to determine the added value of the miRNAs over clinical parameters alone, and increase the generalizability of estimates across cohorts -, and execute a technical replication and validation.

## 2) *By using messenger RNA (mRNA)?*

At least twenty microarray studies on mRNA have aimed to predict response in RA, but so far, the high predictive abilities seem not reproducible in new cohorts. Possibly this irreproducibility is related to the technical drawbacks of microarrays. Therefore, we conducted two separate studies using mRNA to predict response to adalimumab (ADA) and etanercept (ETN) by employing RNA-sequencing (RNA-seq), which is considered to be a more reliable technique.

In **chapter 5** RNA-seq of PBMCs was performed and SEMA6B and GPR15 were selected as promising transcripts to distinguish responders and non-responders. Although these mRNAs were (partially) replicated in new patients, the expression of both genes was strongly associated with smoking, which likely explained the observed differences. However, as both genes play important roles in the pathophysiology of RA and smokers generally experience a reduced treatment success, there is a possibility that SEMA6B and GPR15 are involved in the mechanisms that eventually lead to this reduced therapeutic effect in smokers. In these mechanisms, upregulation of GPR15 and SEMA6B may induce T-cell involvement and synovial vascular proliferation respectively, leading to a more severe and treatment refractory RA. Additionally, we established that mainly the cytokine-related and general immunologic pathways are differentially expressed in responders and non-responders at baseline, which could be explained by the hypothesis in which all bDMARDs eventually target cytokine related pathways.

In the second study employing RNA-seq, **chapter 6**, mRNA of (CD4+) Th-cells was investigated for possible markers that predicted response to ADA or ETN. Across the three patient cohorts, of which one measured by microarray and two by RNA-seq (at two different depths), there was no overlap in predictive transcripts. Although it is likely that technical variability within the microarray plays a role in the different results between the microarray and RNA-seq cohorts, clinical heterogeneity in patient sub-populations may have led to highly variable T-cell transcriptomic profiles, which influences predictor selection and complicates validation. If so, this again stresses the need for the incorporation of clinical parameters in the model that explain the largest heterogeneity, to correct for their influence on the transcriptomic profile and increase the generalizability of results.

### *3) By using proteomics?*

In **chapter 7** we showed that a combination of proteins and clinical parameters predicted response to TNFi treatment slightly better than clinical parameters alone. However, we were not able to replicate these findings in a second cohort of patients, which was possibly related to the large differences in protein concentrations at baseline between the two cohorts. In eight previous studies addressing the same topic, CCL2 showed potential in two studies [1, 2]. This protein was also part of our proteomic platform, yet was not selected as predictive biomarker, and additionally was not identified in two studies employing mass spectrometry [3, 4]. Therefore, CCL2 also has to be considered as not replicated on several occasions, of which once in a targeted platform, indicating there are no robust proteomic predictors to date.

### *4) By using metabolomics?*

Then in **chapter 8**, we found a combination of four metabolites that predicted response to TNFi therapy over clinical parameters alone, which was partially corroborated by internal validation, yet not validated in a separate patient cohort. Besides, it was noticed that the metabolome of RA patients was substantially influenced by clinical parameters, which will complicate any validation and possible application. In particular, glucocorticoids (GCs) influenced the metabolome, which is why we investigated their influence on the lipid

profile in **chapter 9**. Differences in metabolite profile between GC-users and non-users were observed, with increased lysophosphatidylcholines and lysophosphatidylethanolamines, mainly in female GC-users. If replicated, these results have the potential of explaining (some of) the increased cardiovascular risks in GC-users with RA.

*5) By using genomics (in GC response)?*

In **chapter 10**, genetics were employed to explore the relation between polymorphisms in the multidrug receptor 1 (MDR1) gene and response to GCs. Carriage of wild type alleles for any of the three investigated polymorphisms was associated with response, yet only to (high dose) GC pulse therapy. This suggests that wild-type carrying patients might benefit from MDR1 inhibition or administration of glucocorticoid analogues that are no MDR1-substrates. However, these results were obtained in only a small sample set and have to be validated.

Although genomics was not one of the “omics” investigated for possible predictors of response to TNFi treatment in BiOCURA data, thus far other publications have not revealed valuable predictors [5, 6].

**GENERAL DISCUSSION: ARE WE LOST IN VALIDATION?**

In the last decades, many studies have addressed the topic of prediction of response to bDMARDs in RA, and most often (a set of) predictors with potential for clinical practice were proposed. These studies usually state that the only necessary step before fully adopting these predictors, is to carry out a verification study in a separate cohort. This demand is justified, as validation is a key procedure to ensure biologically meaningful and reproducible results [7-10]. Unfortunately, the majority of the results are not being replicated in the same study or ever in a separate validation study. Due to this process of generating unverified results that are not actively validated or contradicted in other cohorts, unusable data is systematically introduced in the field of prediction, without any distinction between applicable and exploratory data. This large heterogeneity in results across studies and lack of validation is also acknowledged in most reviews on this topic [5,

6, 11-22], yet it is fair to assume that the combination of reporting- and publication bias contributes and maintains this process [23, 24]. Each study showing no usable predictors after validation, has to compete with all (unvalidated) positive results already published. It is therefore rewarding for prediction studies to stop after biomarker discovery and only emphasise on these promising exploratory results in the manuscript. Moreover, in a “publish or perish” academic atmosphere, researchers are awarded the most when seeking the easiest and quickest—rather than the best—ways forward [25]. These incentives contribute to the uncertainty if any of these proposed predictors are more than solely false-positive results, and to (publically funded) research waste [23, 25, 26]. Meta-analyses by definition do not exclude negative results from included studies and are thus an answer to overcome reporting bias [24, 26]. They will account for biomarkers that were predictive in one cohort, but not in other studies that incorporated the same marker in its used (limited) platform, such as in a proteomic panel. Still, meta-analyses up till now have not provided robust predictors [27-35], other than a slight increased chance on RTX response for seropositive patients [36]. There is no doubt that improvement in the design of prognostic studies is needed [22, 37]. As an example for initiatives in this issue, TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis, [38]) was started to improve prognostic and diagnostic studies worldwide (<http://www.tripod-statement.org>). More and more journals require a form of validation of results which will probably increase the quality of prognostics studies in the future, although a change in priorities of the academic field is required as well [23, 26]. Above all, external replication in multiple cohorts will always ascertain the trustability of any predictor; thereby honoring the credo of prognostic research: “replication, replication, replication”.

Regardless of the attempts to validate the results from the discovery phase ourselves, this thesis was not able to provide predictors for the prediction of response to therapy in RA, ready to use in clinical practice. This does not mean that the possibility of prediction of response should be ruled out, as many reasons for the inability to find biologically relevant predictors can co-exist. For example, the influence of clinical parameters and technical/biological variability in biomarkers were already pointed out in some chapters of

this thesis. It is therefore conceivable that an adaption of the strategy enables the field to find reproducible and relevant predictors. Main topics in the design of this new strategy that deserve attention were explored by internal discussion, deduction and expert opinion in literature, and are further discussed below.

### **Possible existence of overall bDMARD refractory RA**

Non-response to TNFi's is not well understood [39]. Many studies have demonstrated that TNF $\alpha$  acts as a central mediator within a complex network of cells and cytokines, and most of these functions of TNF $\alpha$  are thought to be inhibited by TNFi's [39]. It is therefore tempting to speculate that in non-responders, the inhibited pathways are bypassed by other pro-inflammatory pathways. If this were true, switching to a bDMARD with a different mode of action after TNFi failure would increase chances of response [40]. Indeed, in a randomised (but unblinded) study of patients failing to their first TNFi, a better response was observed after switching to a non-TNFi than to another TNFi [41]. However, other observations contradict the presence of a bDMARD specific response, and speak for a more general treatment refractory RA. For example, when patients have a history of one or more failed TNFi's, the response to a non-TNFi decreases [42, 43]. Also, if blocking of a specific pathway (e.g. only the TNF $\alpha$  pathway) is insufficient, combining several bDMARDs would inhibit more pro-inflammatory pathways and should increase the response rates [40]. However, although only limitedly evaluated, the added value of combining bDMARDs is not seen in clinical practice [44, 45]. These latter observations imply that all bDMARDs may eventually mediate their efficacy by interfering with a common final pathway—namely, pro-inflammatory cytokine production [40, 46]. RA patients would then either be sensitive to all bDMARDs that interfere with this final pathway, or to some extent bDMARD-refractory. This hypothesis would explain why we found mainly general immunologic and cytokine-related pathways predictive of response to TNFi therapy in PBMC and CD4+ mRNA. However, if predictors of response would be more general for bDMARD-refractory RA than treatment specific, the question still remains why we have not found (good) general predictors of response to bDMARDs yet.

### **Embrace RA heterogeneity as it allows a fair comparison of biomarkers**

There is a high degree of heterogeneity in clinical characteristics among RA patients, which is considered to be a main cause of heterogeneous results among prediction studies [11, 15, 22]. However, heterogeneity also creates appropriate conditions for the application of personalised medicine [47]. For instance, clinical parameters are used in practice to identify patients with the highest risk of rapid radiographic progression or destructive RA disease (in ACR [48, 49] and EULAR [50] guidelines). These general prognostic factors include high disease activity, autoantibody positivity (rheumatoid factor and/or antibodies to citrullinated proteins), the early presence of joint damage (bone erosions), functional limitation (e.g., health assessment questionnaire (HAQ)) and extra-articular disease [51-62]. Prognostic factors are applicable to both treated and untreated patients, and are not to be confused with predictive factors able to distinguish responders from non-responders to a certain treatment [63, 64]. Actually, in order to ascertain the biomarkers are not general prognostic factors but treatment specific, control patients must be included in prediction studies as well [63], e.g. in the discovery phase (as exemplified by Krintel et al.[65]) or in a second validation study. General prognostic factors do not preclude the existence of bDMARD-specific predictors (or refractory RA-specific) and are valuable as they predict the general outcome risk. Prognostic risk models may therefore be considered the basis on which the added predictive value of (treatment specific) biomarkers can be demonstrated.

There is, however, a second and more negative aspect of the clinical heterogeneity that affects prediction studies. In most of the chapters in this thesis, a large effect of clinical parameters on biomarkers was found. Because of this influence, and the fact that clinical characteristics of patients among and between the discovery and validation cohort are heterogeneous, the estimates of (a combination of) biomarkers selected in one cohort will not be easily generalizable to a second cohort [66-70]. Besides, the more subtle effects of relevant predictors might be clouded by heterogeneity and subsequently result in a wrongful rejection [38]. So, although confounding is not traditionally taken into account in prognostic research, other factors can still affect the precision in estimates and validity of the predictors, and increase the risk of bias [70-72]. Incorporation of (a standard set of) clinical parameters in prediction models to “correct” the association between predictor

and outcome, might resolve these issues. Clinical characteristics of relevance may include general prognostic factors, inflammatory marker levels, concomitant treatments, disease duration, comorbidities, the time interval between the last administration of the previous csDMARD/bDMARD before blood was drawn (i.e. washout period), and other factors we do not know yet. The drawback of incorporating (some of) these parameters in biomarker selection and validation, is the necessity of bigger cohorts of patients to account for prediction models with many more variables. Still, these correctional factors are expected to result in a reduced risk of missing biomarkers, the identification of predictors of response to treatment with an added value over general prognostic factors (yet only treatment specific when controls present), and more reliable estimates of predictors resulting in a better generalizability of results.

### **Improvement in prediction modelling and quality of high throughput data**

As for the large volumes of complex data that high-throughput techniques produce, the statistical analyses lie outside the scope of conventional medical and epidemiological practice [73, 74]. Although this thesis investigated the application of multivariable models and supervised clustering techniques, a systems biology approach is theoretically better suited to exploit the full potential of the data, and its application has therefore been encouraged in prediction of response in RA [15, 64, 75]. In systems biology, datasets of high-throughput analyses are combined with advanced mathematical and computational techniques, to screen and analyse the overall genomic, transcriptomic, proteomic and metabolomics state of a cell population. Systems pharmacology, as a subspecialty, focusses on interactions between disease pathophysiology and drug action to obtain mechanistic information on drug responses [73, 76]. Attempts have recently been made in prediction of TNFi response, to integrate transcriptomics (PBMC mRNA), genetics, proteomics, flow cytometry and clinical data derived from the Swedish COMBINE databank, into one model [77]. Predictors proposed in literature were tested for their predictive ability and the best single performing predictors were combined into a model, which showed a good performance. Although biomarker discovery itself was not based on a systems approach and validation was lacking, these initiatives show that RA research is already experimenting with a systems thinking in subclassification of patients.

A second factor that should be taken into account, is that high throughput techniques produce a large amount of data on a relatively small number of biological samples, making the presence of technical variability or noise detrimental to the analyses [11, 78]. Moreover, technical variability can skew the results towards significant differences that are biologically less relevant, leading to the selection of the wrong targets, or obscure prediction of good biomarkers. On some occasions in this thesis the presence of technical noise could be visualised, yet the full extent and influence it may have on the results cannot be estimated. Research should therefore always aim to maximally adjust for noise, by choosing the best profiling technique, standardising procedures and if possible outsourcing to experts, applying appropriate methods for quality control and excluding technical outliers. Regardless of maximal adjustments, noise inherent to the used technique will still remain. However, high-throughput techniques evolve and improve constantly, so technical variability will eventually become less of a problem [79].

#### **High susceptibility in misclassification of response: need for better outcome measures**

A rather obvious prerequisite for research that aims to find predictors of treatment response, is that the response is caused by the pharmacological and pharmacodynamical characteristics of the treatment [63]. However, when the measured response is not a pure pharmacological outcome but contaminated by measurement error and/or mixed with other clinical parameters, prediction will fail to fully succeed. Several examples of misclassification of response are discussed below.

First, a non-response can either be an initial or primary non-response, or a loss of response. An initial non-response, as discussed before, may be caused by a bypass of the pro-inflammatory disease mechanisms next to the (insufficient) pathway blockade by the treatment. In contrast, a loss of response may rather be related to e.g. immunogenicity of the drug, as this takes time to fully develop [80]. In the process of immunogenicity, antidrug-antibodies are created and lead to a decrease in drug concentration and ultimately to an ineffective control of RA disease [81, 82]. Apart from infliximab, which is relatively immunogenic, other TNFi's (adalimumab, etanercept, certolizumab pegol, golimumab) and non-TNFi's (rituximab, tocilizumab, and abatacept) have been shown to elicit such anti-drug antibodies [31, 39, 83, 84]. Methotrexate is known to decrease the

immunogenicity and increases drug survival, which is therefore advised as concomitant treatment to a bDMARD [39, 50, 85, 86]. Concluding, inefficacy on the long term can be a primary and/or secondary failure; each by a different mechanism and therefore likely with its own predictors. So, although sustained remission is the ultimate goal, a longer time-interval might not constitute the most “pure” measurement of response.

A second factor that might induce misclassification of response, is the dependency on a substitute measurement tool for RA disease activity, for example the disease activity based on 28-joint count (DAS28) [87]. Response in BiOCURA was measured using the EULAR response criteria, that are based on the change in DAS28 compared with baseline and absolute DAS28 at that time [88]. Although the DAS28 has been extensively validated and is a widely used instrument in clinical trials, it has its drawbacks. Most relevant for prediction purposes, it has been claimed to be a rather instable monitoring instrument in RA patients with stable disease [89, 90]. This means that “random” fluctuations in disease activity occur, that do not necessarily have a clinical relevance. These fluctuations may lead to a more arbitrary classification of responders and non-responders. Additionally, the DAS28 (as well as CDAI [91] SDAI [92], and ACR-criteria [93],) has the disadvantage that it also omits the assessment of some less regularly affected joints, such as those in the feet. The majority of the patients in BiOCURA with a tender/swollen joint score of zero at baseline had involvement of the feet that prompted the prescription for a TNFi. As a result, these patients were (mis)classified as having a low baseline disease activity, thereby strongly limiting the ability to achieve a EULAR response where a true pharmacological response was as likely to occur. A more stable and objective biochemical monitoring device for RA disease activity would thus be a welcome addition for research purposes. With this goal, a multi-biomarker disease activity (MBDA) test for RA was developed, Vectra™ DA (Crescendo Bioscience) [94], and validated in several external cohorts [95-97]. Baseline MBDA has already been shown to be significantly prognostic for radiographic progression in one study [98], but not in a second study [95]. One question remaining, is whether RA is not a too variable disease to ever allow only one (objective) measurement of disease activity to classify a patient as responder or non-responder: during follow-up visits in BiOCURA, some patients declared that if their response would

have been measured one day earlier results would probably have been completely different.

Lastly, lack of adherence to therapy could have led to misclassification of response as well. Non-compliance is known to exist and influences RA treatment outcomes [99], and thus may have caused some of the observed non-responses in BiOCURA. However, identification of non-compliant patients is challenging and it is therefore difficult to account for in prediction studies. Still, patient education on benefits and risks of therapies and a good communication between physician and patient can generally promote adherence [100], and therefore deserves attention in all studies and clinical practice.

On a final note, maybe we should not be too focused on classifying strictly based on disease activity. It has long been thought that prevention of radiographic progression was achieved through the eradication of synovial inflammation, however, there is evidence that a complete reduction in disease activity is not instrumental in patients on bDMARDs to prevent radiographic progression [101-104]. Possibly, dDMARDs exert sufficient impact on disease processes leading to joint destruction, yet insufficient impact on processes associated with signs and symptoms [101]. If this hypothesis proves to be right, patients with high disease activity are perhaps less at risk for structural damage as feared, and the patient's perspective of response may be the only remaining and relevant outcome to consider. Herein, the most important aspects are pain management and functional improvement, and will include also psychological factors, such as depression and anxiety, that are known to negatively affect response [100].

### **Considerations on other relevant sources of predictors**

It is possible that the biologically most meaningful predictors for RA are not found in the most accessible compartments, such as blood, saliva and urine. Since the disease manifests primarily in the joints, biomarkers measured in synovial tissue might give clearer results [22]. However, (the limited amount of) previous prediction studies on such materials have not resulted in clinically usable predictors yet [22, 105], and the collection of synovial tissue in clinical practice is complicated by feasibility issues [17, 22]. Another issue regarding the choice of the most appropriate compartment, might be the measurement of biomarkers deriving from multiple cell populations, whereas single cell

populations might provide clearer results [17]. On the other hand, we noticed no clear effects of PBMC cell composition and miRNAs or mRNAs levels, and manipulation of cells during isolation procedures may have significant impact on the activity status of such cells, resulting in the induction of extra noise.

Generally, individualised medicine frequently does not consider the exposome (i.e., autoimmune ecology), which characterises both exogenous and endogenous exposures [75]. In BiOCURA only the toxicological influences smoking and alcohol use were taken into account, whereas many more can be of influence including smog inhalation, work exposures, and factors that are frequently believed to have a role in altering RA susceptibility, such as low socioeconomic status, periodontal disease and the gut microbiome [46]. Some of these factors might be important and can be considered to include in future studies.

### **Prediction may be more difficult than foreseen**

The possibility exists that prediction of response is too difficult to become a clinical tool. Not in the least because we and others have not succeeded, but the biology of RA might be too complex and variable to get hold of and/or patients may respond very unpredictable to unknown parameters like the exposome. As Senn has described, studies claiming to observe e.g. 70% response and 30% non-response to treatment, are also consistent with the alternative explanation that 100% of patients respond 70% of the time [106]. Moreover, even with all developing new techniques and increasing knowledge, we remain to be lost in the chaos that eventually lead to the (more easily) observed outcome. There are many parallel circumstances, such as the example in Michael Crichton's *Jurassic Park* (1991), when hitting a ball on a pool table and calculating where it goes. Yes, using Newton's laws you can probably predict where it will go, possibly even for the next few hours. However, as it turns out in practice you can only predict a few seconds into the future, because almost immediately very small effects – imperfections in the surface of the ball, tiny indentations in the wood of the table – start to make a difference. They will eventually even overpower the most careful calculations. Another example, not unrelated in the context of this thesis, is the smoke that rises from a cigarette. The smoke first goes straight up and then spreads out like a diffuse, wavy layer. However, even the best

supercomputers, taken measurements in the most carefully controlled environments, cannot exactly predict the shape of the smoke cloud [107]. By experimenting with study designs that (strongly) simplify the true situation, we might be able to overcome some of the complexity. For example, the bDMARD response can be investigated *in vitro* before starting the bDMARD treatment, possibly by extraction of patient cells (whole blood or synovial tissue to cover a relatively large part of the disease process) and expose these to the treatment. Or leave prediction altogether, and investigate the short term bDMARD response *in vivo* as a pilot study, only looking at the biochemical changes within days after start of bDMARD treatment as an indicator for the expected clinical response. Although these designs are complicated by many factors that are not addressed here, exploring other options might unravel the many parallel processes.

### **Future**

As currently prediction studies have not revealed reproducible and relevant predictors of response to treatment, an adaption of the strategy might be necessary to increase chances of finding them. Changes in design and concept that could facilitate the optimization of prediction studies are provided in **box 1**.

**Box 1:** Suggestions for a change in strategy that might increase chances of finding reproducible predictors of response to therapy in RA.

General efforts to increase the scientific value of studies and keep an overview:

- Following guidelines to ascertain high quality of studies, both generally for prediction studies (e.g. TRIPOD [38]), and specifically for quality of “omics” (e.g. MIAME for microarray data [108]).
- Validate findings in a separate cohort, and in case that markers cannot be validated, publish these negative results.
- Promote acceptance of only validated results. One idea is to code all prediction studies with one of two marks: “replicated” or “unchallenged”[109].
- Facilitate meta-analysis of individual participant data from prognostic factor studies through data sharing initiatives [63].

Source of biomarkers and measurement:

- Consider both mixed- and single cell populations.
- Consider the exposome.
- Choose a robust profiling technique, standardise methods, perform quality control, and exclude technical outliers.

Classification of response:

- Take a short time-interval to reduce inclusion of non-response due to immunogenicity.
- Maximise adherence to therapy by maintaining a good communication.
- On the long term, improve measurement of disease activity and definition of response:
  - o pure biochemical response measurements.
  - o incorporating the minimal level of disease activity resulting in radiographic damage.

Embracing RA heterogeneity:

- Correct biomarkers for possible influencing clinical parameters, to increase estimates of predictive value and generalizability
- Compare the added value of (expensive) biomarkers over general prognostic predictors.
- Before fully accepting the predictor, compare biomarkers between treated and untreated patients to distinct between general prognostic and treatment specific predictors.

Other designs to comprehend (some of) the complexity:

- Explore the actual systems biology/pharmacology approaches by integrating different omics platforms.
- Experiment with short term in vitro response/short term (system) biochemical in vivo response.

Lastly, in the field of RA there is an overall agreement that changing RA from a manageable into a curable disease entity will eventually require identification of the exact etiologic factors and initiating pathways [110]. With prediction it might be alike: understanding of the pathogenesis and pharmacological effects of therapy might be needed for successful application of individualised medicine [111, 112]. This means that instead of the current 'shotgun approach' in which many biomarkers at once are measured and some might hit the target, energy and funding might possibly better be spent in understanding the pathophysiology and pharmacological response mechanisms. Using this knowledge, hypotheses can be made that can be specifically tested, while also taking into account all (biologically possible) influencing factors. Although this might seem a step backwards, it can actually help us forward.

## **CONCLUSION**

Despite all efforts in this thesis and by others before, usable predictors of response to therapy in RA are still lacking. Although leading researchers in rheumatology firmly believe that predictors for treatment response in RA will be available someday to use [46], this will probably necessitate a shift in the design of prediction studies, while taking into account the issues discussed in this chapter.

## REFERENCES

1. Fabre S, Dupuy AM, Dossat N, Guisset C, Cohen JD, Cristol JP et al. Protein biochip array technology for cytokine profiling predicts etanercept responsiveness in rheumatoid arthritis. *Clin Exp Immunol* 2008;153 2: 188-95.
2. Hueber W, Tomooka BH, Batliwalla F, Li W, Monach PA, Tibshirani RJ et al. Blood autoantibody and cytokine profiles predict response to anti-tumor necrosis factor therapy in rheumatoid arthritis. *Arthritis Res Ther* 2009;11 3: R76.
3. Ortea I, Roschitzki B, Ovalles JG, Longo JL, de IT, I, Gonzalez I et al. Discovery of serum proteomic biomarkers for prediction of response to infliximab (a monoclonal anti-TNF antibody) treatment in rheumatoid arthritis: an exploratory analysis. *J Proteomics* 2012;77: 372-82.
4. Trocme C, Marotte H, Baillet A, Pallot-Prades B, Garin J, Grange L et al. Apolipoprotein A-I and platelet factor 4 are biomarkers for infliximab response in rheumatoid arthritis. *Ann Rheum Dis* 2009;68 8: 1328-33.
5. Goulielmos GN, Zervou MI, Myrthianou E, Burska A, Niewold TB, Ponchel F. Genetic data: The new challenge of personalized medicine, insights for rheumatoid arthritis patients. *Gene* 2016;583 2: 90-101.
6. Xie X, Zhang D, Chen JW, Tian J, Ling GH, Li F. Pharmacogenomics of biological treatment in rheumatoid arthritis. *Expert Opin Biol Ther* 2014;14 2: 157-64.
7. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013;10 2: e1001381.
8. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ* 2009;338: b375.
9. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009;338: b605.
10. Horvath AR, Kis E, Dobos E. Guidelines for the use of biomarkers: principles, processes and practical considerations. *Scand J Clin Lab Invest Suppl* 2010;242: 109-16.
11. Burska AN, Roget K, Blits M, Soto Gomez L, van de Loo F, Hazelwood LD et al. Gene expression analysis in RA: towards personalized medicine. *Pharmacogenomics J* 2014;14 2: 93-106.
12. Daïen CI, Morel J. Predictive factors of response to biological disease modifying antirheumatic drugs: towards personalized medicine. *Mediators Inflamm* 2014;2014: 386148.
13. Isaacs JD, Ferraccioli G. The need for personalised medicine for rheumatoid arthritis. *Ann Rheum Dis* 2011;70 1: 4-7.
14. Katchamart W, Johnson S, Lin HJ, Phumethum V, Salliot C, Bombardier C. Predictors for remission in rheumatoid arthritis patients: A systematic review. *Arthritis Care Res (Hoboken)* 2010;62 8: 1128-43.
15. Marotte H, Miossec P. Biomarkers for prediction of TNFalpha blockers response in rheumatoid arthritis. *Joint Bone Spine* 2010;77 4: 297-305.
16. Perricone C, Ceccarelli F, Valesini G. An overview on the genetic of rheumatoid arthritis: a never-ending story. *Autoimmun Rev* 2011;10 10: 599-608.
17. Smith SL, Plant D, Eyre S, Barton A. The potential use of expression profiling: implications for predicting treatment response in rheumatoid arthritis. *Ann Rheum Dis* 2013.
18. van den Broek M, Visser K, Allaart CF, Huizinga TW. Personalized medicine: predicting responses to therapy in patients with RA. *Curr Opin Pharmacol* 2013.
19. Simsek I. Predictors of response to TNF inhibitors in rheumatoid arthritis - do we have new tools for personalized medicine? *Bull NYU Hosp Jt Dis* 2012;70 3: 187-90.
20. van Baarsen LG, Bos CL, TC vdPK, Verweij CL. Transcription profiling of rheumatic diseases. *Arthritis Res Ther* 2009;11 1: 207.
21. Emery P, Dornier T. Optimising treatment in rheumatoid arthritis: a review of potential biological markers of response. *Ann Rheum Dis* 2011;70 12: 2063-70.
22. Salgado E, Maneiro JR, Gomez-Reino JJ. Predictors of response to TNF antagonists. *Curr Pharm Des* 2015;21 2: 221-32.
23. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383 9912: 166-75.
24. Kyzas PA, Loizou KT, Ioannidis JP. Selective reporting biases in cancer prognostic factor

- studies. *J Natl Cancer Inst* 2005;97 14: 1043-55.
25. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP et al. Biomedical research: increasing value, reducing waste. *Lancet* 2014;383 9912: 101-4.
  26. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gulmezoglu AM et al. How to increase value and reduce waste when research priorities are set. *Lancet* 2014;383 9912: 156-65.
  27. Sieberts SK, Zhu F, Garcia-Garcia J, Stahl E, Pratap A, Pandey G et al. Crowdsourced assessment of common genetic contribution to predicting anti-TNF treatment response in rheumatoid arthritis. *Nat Commun* 2016;7: 12460.
  28. Lee YH, Ji JD, Bae SC, Song GG. Associations between tumor necrosis factor-alpha (TNF-alpha) -308 and -238 G/A polymorphisms and shared epitope status and responsiveness to TNF-alpha blockers in rheumatoid arthritis: a metaanalysis update. *J Rheumatol* 2010;37 4: 740-6.
  29. Lee YH, Bae SC, Song GG. Meta-analysis of gene expression profiles to predict response to biologic agents in rheumatoid arthritis. *Clin Rheumatol* 2014;33 6: 775-82.
  30. Mandema JW, Salinger DH, Baumgartner SW, Gibbs MA. A dose-response meta-analysis for quantifying relative efficacy of biologics in rheumatoid arthritis. *Clin Pharmacol Ther* 2011;90 6: 828-35.
  31. Maneiro JR, Salgado E, Gomez-Reino JJ. Immunogenicity of monoclonal antibodies against tumor necrosis factor used in chronic immune-mediated inflammatory conditions: systematic review and meta-analysis. *JAMA Intern Med* 2013;173 15: 1416-28.
  32. Maneiro RJ, Salgado E, Carmona L, Gomez-Reino JJ. Rheumatoid factor as predictor of response to abatacept, rituximab and tocilizumab in rheumatoid arthritis: Systematic review and meta-analysis. *Semin Arthritis Rheum* 2013;43 1: 9-17.
  33. Pavy S, Toonen EJ, Miceli-Richard C, Barrera P, van Riel PL, Criswell LA et al. Tumour necrosis factor alpha -308G->A polymorphism is not associated with response to TNFalpha blockers in Caucasian patients with rheumatoid arthritis: systematic review and meta-analysis. *Ann Rheum Dis* 2010;69 6: 1022-8.
  34. Salgado E, Maneiro JR, Carmona L, Gomez-Reino J. Rheumatoid factor and response to TNF antagonists in rheumatoid arthritis: systematic review and meta-analysis of observational studies. *Joint Bone Spine* 2014;81 1: 41-50.
  35. Zeng Z, Duan Z, Zhang T, Wang S, Li G, Gao J et al. Association between tumor necrosis factor-alpha (TNF-alpha) promoter -308 G/A and response to TNF-alpha blockers in rheumatoid arthritis: a meta-analysis. *Mod Rheumatol* 2013;23 3: 489-95.
  36. Isaacs JD, Cohen SB, Emery P, Tak PP, Wang J, Lei G et al. Effect of baseline rheumatoid factor and anticitrullinated peptide antibody serotype on rituximab clinical response: a meta-analysis. *Ann Rheum Dis* 2013;72 3: 329-36.
  37. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med* 2013;10 2: e1001380.
  38. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162 1: W1-73.
  39. Tracey D, Klareskog L, Sasso EH, Salfeld JG, Tak PP. Tumor necrosis factor antagonist mechanisms of action: a comprehensive review. *Pharmacol Ther* 2008;117 2: 244-79.
  40. Smolen JS, Aletaha D. Forget personalised medicine and focus on abating disease activity. *Ann Rheum Dis* 2013;72 1: 3-6.
  41. Gottenberg JE, Brocq O, Perdriger A, Lassoued S, Berthelot JM, Wendling D et al. Non-TNF-Targeted Biologic vs a Second Anti-TNF Drug to Treat Rheumatoid Arthritis in Patients With Insufficient Response to a First Anti-TNF Drug: A Randomized Clinical Trial. *JAMA* 2016;316 11: 1172-80.
  42. Schiff M, Pritchard C, Huffstutter JE, Rodriguez-Valverde V, Durez P, Zhou X et al. The 6-month safety and efficacy of abatacept in patients with rheumatoid arthritis who underwent a washout after anti-tumour necrosis factor therapy or were directly switched to abatacept: the ARRIVE trial. *Ann Rheum Dis* 2009;68 11: 1708-14.
  43. Harrold LR, Reed GW, Shewade A, Magner R, Saunders KC, John A et al. Effectiveness of Rituximab for the Treatment of Rheumatoid Arthritis in Patients with Prior Exposure to Anti-TNF: Results from the CORRONA Registry. *J Rheumatol* 2015;42 7: 1090-8.
  44. Weinblatt M, Schiff M, Goldman A, Kremer J, Luggen M, Li T et al. Selective costimulation modulation using abatacept in patients with active rheumatoid arthritis

- while receiving etanercept: a randomised clinical trial. *Ann Rheum Dis* 2007;66 2: 228-34.
45. Genovese MC, Cohen S, Moreland L, Lium D, Robbins S, Newmark R et al. Combination therapy with etanercept and anakinra in the treatment of patients with rheumatoid arthritis who have been treated unsuccessfully with methotrexate. *Arthritis Rheum* 2004;50 5: 1412-9.
  46. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet* 2016.
  47. Miossec P, Verweij CL, Klareskog L, Pitzalis C, Barton A, Lekkerkerker F et al. Biomarkers and personalised medicine in rheumatoid arthritis: a proposal for interactions between academia, industry and regulatory bodies. *Ann Rheum Dis* 2011;70 10: 1713-8.
  48. Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol* 2016;68 1: 1-26.
  49. Singh JA, Furst DE, Bharat A, Curtis JR, Kavanaugh AF, Kremer JM et al. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2012;64 5: 625-39.
  50. Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73 3: 492-509.
  51. Visser K, Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Roday HK, Seys PE, Kerstens PJ et al. A matrix risk model for the prediction of rapid radiographic progression in patients with rheumatoid arthritis receiving different dynamic treatment strategies: post hoc analyses from the BeSt study. *Ann Rheum Dis* 2010;69 7: 1333-7.
  52. Vastesaeger N, Xu S, Aletaha D, St Clair EW, Smolen JS. A pilot risk model for the prediction of rapid radiographic progression in rheumatoid arthritis. *Rheumatology (Oxford)* 2009;48 9: 1114-21.
  53. van der Heijde DM, van Riel PL, van Leeuwen MA, van 't Hof MA, van Rijswijk MH, van de Putte LB. Prognostic factors for radiographic damage and physical disability in early rheumatoid arthritis. A prospective follow-up study of 147 patients. *Br J Rheumatol* 1992;31 8: 519-25.
  54. Scott DL, Grindulis KA, Struthers GR, Coulton BL, Popert AJ, Bacon PA. Progression of radiological changes in rheumatoid arthritis. *Ann Rheum Dis* 1984;43 1: 8-17.
  55. Fries JF, Bloch DA, Sharp JT, McShane DJ, Spitz P, Bluhm GB et al. Assessment of radiologic progression in rheumatoid arthritis. A randomized, controlled trial. *Arthritis Rheum* 1986;29 1: 1-9.
  56. Krabben A, Huizinga TW, Mil AH. Biomarkers for radiographic progression in rheumatoid arthritis. *Curr Pharm Des* 2015;21 2: 147-69.
  57. Agrawal S, Misra R, Aggarwal A. Autoantibodies in rheumatoid arthritis: association with severity of disease in established RA. *Clin Rheumatol* 2007;26 2: 201-4.
  58. Ates A, Kinikli G, Turgay M, Akay G, Tokgoz G. Effects of rheumatoid factor isotypes on disease activity and severity in patients with rheumatoid arthritis: a comparative study. *Clin Rheumatol* 2007;26 4: 538-45.
  59. Berglin E, Johansson T, Sundin U, Jidell E, Wadell G, Hallmans G et al. Radiological outcome in rheumatoid arthritis is predicted by presence of antibodies against cyclic citrullinated peptide before and at disease onset, and by IgA-RF at disease onset. *Ann Rheum Dis* 2006;65 4: 453-8.
  60. Meyer O, Labarre C, Dougados M, Goupille P, Cantagrel A, Dubois A et al. Anticitrullinated protein/peptide antibody assays in early rheumatoid arthritis for predicting five year radiographic damage. *Ann Rheum Dis* 2003;62 2: 120-6.
  61. Vencovsky J, Machacek S, Sedova L, Kafkova J, Gatterova J, Pesakova V et al. Autoantibodies can be prognostic markers of an erosive disease in early rheumatoid arthritis. *Ann Rheum Dis* 2003;62 5: 427-30.
  62. Brook A, Corbett M. Radiographic changes in early rheumatoid disease. *Ann Rheum Dis* 1977;36 1: 71-3.
  63. Hingorani AD, Windt DA, Riley RD, Abrams K, Moons KG, Steyerberg EW et al. Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ* 2013;346: e5793.
  64. MacIsaac KD, Wang IM, Menetski J, Roberts C. Genomic and systems approaches to translational biomarker discovery in immunological diseases. *Drug Discov Today* 2014;19 2: 133-9.
  65. Krintel SB, Dehlerdorff C, Hetland ML, Horslev-Petersen K, Andersen KK, Junker P et al. Prediction of treatment response to adalimumab: a double-blind placebo-

- controlled study of circulating microRNA in patients with early rheumatoid arthritis. *Pharmacogenomics J* 2015.
66. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 2009;338: b606.
  67. Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol* 2008;61 1: 76-86.
  68. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Stat Med* 2004;23 16: 2567-86.
  69. van Houwelingen HC. Validation, calibration, revision and combination of prognostic survival models. *Stat Med* 2000;19 24: 3401-15.
  70. Steyerberg EW. *Clinical Prediction Models*. Publisher: Springer Science 2009: ISBN: 978-0-387-77243-1.
  71. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;158 4: 280-6.
  72. Huguet A, Hayden JA, Stinson J, McGrath PJ, Chambers CT, Tougas ME et al. Judging the quality of evidence in reviews of prognostic factor research: adapting the GRADE framework. *Syst Rev* 2013;2: 71.
  73. Jain K. *Textbook of Personalized Medicine*. Publisher: Springer Science 2016: ISBN: 978-1-4939-2552-0.
  74. Ioannidis JP. Expectations, validity, and reality in omics. *J Clin Epidemiol* 2010;63 9: 945-9.
  75. Anaya JM, Duarte-Rey C, Sarmiento-Monroy JC, Bardey D, Castiblanco J, Rojas-Villarraga A. Personalized medicine. Closing the gap between knowledge and clinical practice. *Autoimmun Rev* 2016;15 8: 833-42.
  76. Wist AD, Berger SI, Iyengar R. Systems pharmacology and genome medicine: a future perspective. *Genome Med* 2009;1 1: 11.
  77. Folkersen L, Brynedal B, Diaz-Gallo LM, Ramskold D, Shchetynsky K, Westerlind H et al. Integration of known DNA, RNA and protein biomarkers provides prediction of anti-TNF response in rheumatoid arthritis: results from the COMBINE study. *Mol Med* 2016;22.
  78. Mayeux R. Biomarkers: potential uses and limitations. *NeuroRx* 2004;1 2: 182-8.
  79. Marsal S, Julia A. Rheumatoid arthritis pharmacogenomics. *Pharmacogenomics* 2010;11 5: 617-9.
  80. Mok CC, Tsai WC, Chen DY, Wei JC. Immunogenicity of anti-TNF biologic agents in the treatment of rheumatoid arthritis. *Expert Opin Biol Ther* 2016;16 2: 201-11.
  81. Wolbink GJ, Voskuyl AE, Lems WF, de GE, Nurmohamed MT, Tak PP et al. Relationship between serum trough infliximab levels, pretreatment C reactive protein levels, and clinical response to infliximab treatment in patients with rheumatoid arthritis. *Ann Rheum Dis* 2005;64 5: 704-7.
  82. Jani M, Chinoy H, Warren RB, Griffiths CE, Plant D, Fu B et al. Clinical utility of random anti-tumor necrosis factor drug-level testing and measurement of antidrug antibodies on the long-term treatment response in rheumatoid arthritis. *Arthritis Rheumatol* 2015;67 8: 2011-9.
  83. Jani M, Isaacs JD, Morgan AW, Wilson AG, Plant D, Hyrich KL et al. High frequency of antidrug antibodies and association of random drug levels with efficacy in certolizumab pegol-treated patients with rheumatoid arthritis: results from the BRAGGSS cohort. *Ann Rheum Dis* 2016.
  84. Kneepkens EL, Plasencia C, Krieckaert CL, Pascual-Salcedo D, van der Kleij D, Nurmohamed MT et al. Golimumab trough levels, antidrug antibodies and clinical response in patients with rheumatoid arthritis treated in daily clinical practice. *Ann Rheum Dis* 2014;73 12: 2217-9.
  85. Garces S, Demengeot J, Benito-Garcia E. The immunogenicity of anti-TNF therapy in immune-mediated inflammatory diseases: a systematic review of the literature with a meta-analysis. *Ann Rheum Dis* 2013;72 12: 1947-55.
  86. Cuppen BV, Jacobs JW, Ter Borg EJ, Marijnissen AC, Bijlsma JW, Lafeber FP et al. Necessity of TNF-alpha inhibitor discontinuation in rheumatoid arthritis is predicted by smoking and number of previously used biological DMARDs. *Clin Exp Rheumatol* 2016.
  87. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38 1: 44-8.

88. van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;41 10: 1845-50.
89. Behrens F, Tony HP, Alten R, Kleinert S, Scharbatke EC, Kohm M et al. Development and validation of a new DAS28-based treatment response criterion for rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2013.
90. Madsen OR. Stability of clinical outcome measures in rheumatoid arthritis patients with stable disease defined on the basis of the EULAR response criteria. *Clin Rheumatol* 2016.
91. Aletaha D, Nell V, Stamm T, Uffmann M, Pflugbeil S, Machold K et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Research and Therapy* 2005;7: R796-R806.
92. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology* 2003;42: 244-57.
93. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;38 6.
94. Centola M, Cavet G, Shen Y, Ramanujan S, Knowlton N, Swan KA et al. Development of a multi-biomarker disease activity test for rheumatoid arthritis. *PLoS One* 2013;8 4: e60635.
95. Bakker MF, Cavet G, Jacobs JW, Bijlsma JW, Haney DJ, Shen Y et al. Performance of a multi-biomarker score measuring rheumatoid arthritis disease activity in the CAMERA tight control study. *Ann Rheum Dis* 2012;71 10: 1692-7.
96. Curtis JR, van der Helm-van Mil AH, Knevel R, Huizinga TW, Haney DJ, Shen Y et al. Validation of a novel multibiomarker test to assess rheumatoid arthritis disease activity. *Arthritis Care Res (Hoboken)* 2012;64 12: 1794-803.
97. Hirata S, Dirven L, Shen Y, Centola M, Cavet G, Lems WF et al. A multi-biomarker score measures rheumatoid arthritis disease activity in the BeSt study. *Rheumatology (Oxford)* 2013;52 7: 1202-7.
98. Hambardzumyan K, Bolce RJ, Saevardottir S, Forslind K, Wallman JK, Cruickshank SE et al. Association of a multibiomarker disease activity score at multiple time-points with radiographic progression in rheumatoid arthritis: results from the SWEFOT trial. *RMD Open* 2016;2 1: e000197.
99. Pascual-Ramos V, Contreras-Yanez I, Villa AR, Cabiedes J, Rull-Gabayet M. Medication persistence over 2 years of follow-up in a cohort of early rheumatoid arthritis patients: associated factors and relationship with disease activity and with disability. *Arthritis Res Ther* 2009;11 1: R26.
100. Santiago T, Geenen R, Jacobs JW, da Silva JA. Psychological factors associated with response to treatment in rheumatoid arthritis. *Curr Pharm Des* 2015;21 2: 257-69.
101. Villeneuve E, Haraoui B. Uncoupling of disease activity and structural damage. Does it matter clinically? *Ann Rheum Dis* 2013;72 1: 1-2.
102. Landewe R, van der Heijde D, Klareskog L, van Vollenhoven R, Fatenejad S. Disconnect between inflammation and joint destruction after treatment with etanercept plus methotrexate: results from the trial of etanercept and methotrexate with radiographic and patient outcomes. *Arthritis Rheum* 2006;54 10: 3119-25.
103. Smolen JS, van der Heijde DM, Keystone EC, van Vollenhoven RF, Goldring MB, Guertel B et al. Association of joint space narrowing with impairment of physical function and work ability in patients with early rheumatoid arthritis: protection beyond disease control by adalimumab plus methotrexate. *Ann Rheum Dis* 2013;72 7: 1156-62.
104. Smolen JS, Han C, Bala M, Maini RN, Kalden JR, van der Heijde D et al. Evidence of radiographic benefit of treatment with infliximab plus methotrexate in rheumatoid arthritis patients who had no clinical improvement: a detailed subanalysis of data from the anti-tumor necrosis factor trial in rheumatoid arthritis with concomitant therapy study. *Arthritis Rheum* 2005;52 4: 1020-30.
105. de Hair MJ, Harty LC, Gerlag DM, Pitzalis C, Veale DJ, Tak PP. Synovial tissue analysis for the discovery of diagnostic and prognostic biomarkers in patients with early arthritis. *J Rheumatol* 2011;38 9: 2068-72.
106. Senn S. Individual response to treatment: is it a valid assumption? *BMJ* 2004;329 7472: 966-8.
107. Bryson B. A short history of nearly everything. Publisher: Black Swan 2003: ISBN: 9780552997041.

108. Edgar R, Barrett T. NCBI GEO standards and services for microarray data. *Nat Biotechnol* 2006;24 12: 1471-2.
109. Ioannidis JP. Why Science Is Not Necessarily Self-Correcting. *Perspect Psychol Sci* 2012;7 6: 645-54.
110. Goronzy JJ, Weyand CM. Developments in the scientific understanding of rheumatoid arthritis. *Arthritis Res Ther* 2009;11 5: 249.
111. Ma Q, Lu AY. Pharmacogenetics, pharmacogenomics, and individualized medicine. *Pharmacol Rev* 2011;63 2: 437-59.
112. Baker M. In biomarkers we trust? *Nat Biotechnol* 2005;23 3: 297-304.





## **Chapter 12**

**Nederlandse samenvatting**

**Dankwoord**

**Curriculum vitae**

**List of publications**

## **Nederlandse samenvatting**

Gewrichtsreuma of reumatoïde artritis (RA) is een auto-immuunziekte die primair leidt tot ontstekingen aan de gewrichten. De exacte ontstaanswijze van RA is nog niet volledig opgehelderd, maar bepaalde erfelijke factoren en blootstellingen zoals roken, tandvleesontstekingen, de darmflora en nog onbekende factoren, leiden uiteindelijk tot een situatie waarin witte bloedcellen het eigen lichaam aanvallen. De RA-patiënt ervaart vooral pijn, zwelling en functieverlies van de aangedane gewrichten, met op de langere termijn kans op irreversibele gewrichtsschade en invaliditeit. Daarnaast treden tal van andere verschijnselen op zoals moeheid, (lichte) koorts, gezwollen lymfeklieren en vorming van onderhuidse knobbels. Omdat RA vooralsnog niet te genezen is, dient de ontsteking veroorzaakt door de witte bloedcellen maximaal gecontroleerd te worden, om zowel goed te kunnen functioneren in het dagelijks leven alsmede lange termijn schade te voorkomen. Voor de behandeling is een breed arsenaal aan ontstekingsremmende middelen beschikbaar, waarvan methotrexaat en prednison de meest gebruikte conventionele therapieën zijn. Bij een nieuw ontdekte reuma wordt er meestal met methotrexaat (al dan niet in combinatie met prednison) gestart gezien de ruime ervaring, bewezen werkzaamheid en lage kosten. De laatste decaden hebben ook enkele zeer geavanceerde middelen hun intrede gedaan, die gericht zijn op specifieke ontstekingsgerelateerde eiwitten, de zogenaamde “biologicaltherapieën” of kortweg “biologicals”. Deze nieuwe middelen hebben de behandeling van RA en andere auto-immuunaandoeningen revolutionair beïnvloed.

Ondanks het brede arsenaal aan traditionele middelen en biologicals is de behandeling van RA uitdagend, aangezien de gekozen therapie niet altijd effectief genoeg de ontsteking remt, ofwel (te) veel bijwerkingen geeft. In deze gevallen wordt vaak overgestapt naar een ander middel met de hoop op een betere uitkomst. Deze ‘trial and error’ benadering is vooralsnog het beste wat mogelijk is, omdat elke patiënt verschillend reageert op de therapie. Echter, omdat de effectiviteit van de therapie pas na enkele maanden goed vast te stellen is, zullen de slecht reagerende patiënten voor langere periode een ongecontroleerde ziekte doormaken en al die tijd kampen met (forse) ontstekingen, met mogelijke restschade tot gevolg. Idealiter zou bij voorbaat al bekend zijn hoe de patiënt op een bepaald middel gaat reageren, om zo de beste behandeling

voor die patiënt te starten. Er moet dan gezocht worden naar verschillende typen RA of patiënt-factoren die maken dat de ene patiënt wel reageert en de andere niet. Een bloedtest op dergelijke factoren geeft dan bijvoorbeeld de kans op slagen weer op een therapie, zodat arts en patiënt een goede afweging voor al dan niet starten kunnen maken. Deze test zou het dagelijks functioneren van patiënten en de lange termijn uitkomsten sterk kunnen verbeteren, met alle persoonlijke en maatschappelijke voordelen van dien. Aangezien de biologicals relatief dure geneesmiddelen zijn, zou een effectiever gebruik (of switch naar traditionele middelen) ook in directe zin kostenbesparend kunnen zijn. Een toekomst waarin predictie van therapierespons mogelijk is, is daarmee zeer wenselijk voor zowel patiënt, arts als maatschappij. In dit proefschrift is gepoogd om met behulp van patiëntkarakteristieken de respons op therapie voor een persoon te voorspellen.

In dit proefschrift werd vooral gebruik gemaakt van data uit het onderzoek “biologicals and outcomes, compared and predicted in the Utrecht region in rheumatoid arthritis” (BiOCURA), waarin RA-patiënten gevolgd zijn vanaf het moment dat zij startten met een biological tot één jaar daarna. Respons werd in BiOCURA gemeten aan de hand van verbetering in pijnlijke en gezwollen gewrichten, alsmede verbetering in de ontstekingswaarden in het bloed en algehele gezondheidsscore aangegeven door de patiënt, de zogenaamde EULAR respons criteria. Omdat van deze patiënten bloed en urine is afgenomen voor start van de therapie, en de effectiviteit geobserveerd is na drie, zes en twaalf maanden, kan gezocht worden naar factoren die vóór start van de therapie voorspellend zijn voor een goede of slechte klinische respons. Van deze losse voorspellers kunnen weer modellen worden gemaakt, waarin een algoritme per individu bepaalt wat de kansen op respons zijn. In het BiOCURA onderzoek werden gegevens en biomaterialen van patiënten verzameld in acht ziekenhuizen die aangesloten zijn bij de Stichting Reumaonderzoek Utrecht (SRU). Er werden geen patiënten uitgesloten voor deelname en de beslissingen tussen arts en patiënt (zoals het toevoegen van andere middelen of keuze tot stoppen met therapie) werden niet beïnvloed. Daarmee is de BiOCURA een geheel observationeel onderzoek dat de klinische praktijk zo zuiver mogelijk weergeeft.

*De volgende vragen werden beantwoord in het proefschrift:*

*Wat weten we van markers voor respons in RA?*

In hoofdstuk 2 werd een literatuuronderzoek verricht naar de uitkomsten van eerdere onderzoeken naar de respons op biologicals. Er werden meer dan vijftig studies gevonden die voorspellende factoren suggereerden te hebben gevonden, zoals klinische patiëntkarakteristieken (leeftijd, geslacht, BMI etc.), genen en eiwitten. Echter, wanneer de uitkomsten tussen studies werden vergeleken, waren er geen universele factoren die de kansen op respons per patiënt in relevante mate beïnvloedden. Bij start van dit onderzoek waren er dus geen relevante voorspellende factoren voor respons bekend.

*Wie stoppen of continueren met therapie?*

In hoofdstuk 3 werd onderzocht welke karakteristieken van patiënten iets zeggen over de kans op het stoppen van therapie, binnen een jaar na start. Patiënten die gestart waren met TNF-alfa blokkerende therapie (TNF-inhibitors / TNFi), die de grootste groep van biologicals omvat, werden hiervoor geselecteerd. Hieruit bleek dat patiënten die meerdere biologicals in het verleden (onsuccesvol) gebruikt hadden en patiënten die rookten, het vaakst de TNFi behandeling stakten. Deze uitkomsten verschaffen interessante informatie over factoren die het behandelingsucces significant beïnvloeden, en roept vooral vragen op over het achterliggende mechanisme. Echter, omdat deze factoren maar een klein deel van de geobserveerde verschillen kunnen verklaren, is het op individueel patiëtniveau niet mogelijk om voorspellende factoren te identificeren en dus niet bruikbaar voor de praktijk. Met andere woorden, de resultaten zijn significant en interessant, maar niet relevant.

In de volgende hoofdstukken werden potentiële voorspellers voor therapierespons in bloed en urine gemeten, zoals DNA, RNA, eiwitten en metaboliëten (zie ook Figure 2 in hoofdstuk 1, verderop wordt hun individuele rol verder uitgelegd). Deze factoren worden overkoepelend ook wel 'biomarkers' genoemd, omdat ze een marker of afspiegeling zijn van biologische processen in het lichaam. Het ziekteproces van RA wordt bepaald door

deze biologische processen, en mogelijk daarmee de respons op therapie ook. De uitdaging ligt in het vinden van de beste (combinatie van) biomarkers in de grote hoeveelheden meetbare markers.

*Kunnen we betere factoren vinden:*

*Met gebruik van microRNA's (miRNA's)?*

In hoofdstuk 4 onderzochten we >750 miRNA's uit het bloed. miRNA's zijn kleine moleculen die invloed kunnen uitoefenen op DNA en RNA, en zo de uiteindelijke eiwitproductie beïnvloeden. In 80 patiënten die gestart waren met één van twee TNFi's, adalimumab (ADA) of etanercept (ETN), werden vier potentieel interessante miRNA's gevonden die respons voorspelden. Echter, in een nieuwe groep van 40 patiënten konden de oorspronkelijke resultaten niet gevalideerd worden. Dit zou kunnen betekenen dat de oorspronkelijke resultaten wellicht toevalsbevindingen waren die geen biologische betekenis hebben. Daarnaast konden we de resultaten niet repliceren van twee eerdere studies die miRNA's gesuggereerd hadden. Mogelijk weken onze studie-patiënten en metingen te veel af van de oorspronkelijke studies, ofwel waren de resultaten uit deze eerdere studies ook toevalsbevindingen: beide studies hadden hun resultaten niet zelf gevalideerd in een nieuwe groep patiënten, wat zeer belangrijk is om het effect van toeval te elimineren bij dit soort grote aantallen gemeten biomarkers in een kleine groep patiënten.

*Met gebruik van messenger RNA (mRNA)?*

Vervolgens werd in hoofdstuk 5 het mRNA van lymfocyten (type witte bloedcellen) gemeten en gebruikt voor voorspelling van respons op ADA en ETN. mRNA's zijn een directe afspiegeling van de genen in het DNA die 'actief' zijn in een specifieke cel, en vormen de boodschap voor het maken van eiwitten. Meer dan 15 eerdere studies op dit onderwerp gebruikten de traditionele microarray techniek voor mRNA-bepaling en vonden geen voorspellers voor respons die herhaald konden worden in ander onderzoek, echter deze techniek kent veel beperkingen. De metingen in het onderzoek in hoofdstuk 5 werden verricht met de nieuwere RNA-sequencing techniek. Aanvankelijk werden in 80

patiënten met goede of slechte respons twee potentieel interessante mRNA's gevonden (uit >60.000 gemeten mRNA's). Deze twee mRNA's werden in dezelfde groep van 80 patiënten ook gemeten met kwantitatieve PCR. Dit is een gevoelige analysetechniek waarmee je een specifiek mRNA kunt meten. Door de bevestiging met de kwantitatieve PCR werd een meetfout door RNA-sequencing uitgesloten. Echter, in een nieuwe set van 40 patiënten gemeten met de verschillende qPCR's, bleken de mRNA's van ADA wél, maar die voor ETA niet opnieuw van waarde. Dit opvallende verschil kon verklaard worden door de verdeling van het aantal rokers tussen de groepen, met hogere aantallen rokers onder de patiënten zonder respons bij ADA, maar niet bij ETN. De twee geselecteerde mRNA's, SEMA6B & GPR15, bleken namelijk een sterke afspiegeling van roken te zijn. Omdat deze mRNA's ook aan RA-ontsteking gerelateerd zijn, zou het kunnen dat roken in de praktijk zijn negatieve effect op RA en effectiviteit van therapie uitoefent via deze mRNA's.

mRNA en voorspelling van ADA en ETN therapierespons werd ook onderzocht in hoofdstuk 6, maar dan van mRNA van T-helper-cellen, een subtype lymfocyt met grote rol in RA. In drie patiëntgroepen werd met behulp van microarray (40 patiënten), en RNA-sequencing (14 en 40 patiënten) potentieel interessante mRNAs geselecteerd en de overlap in gevonden biomarkers onderzocht (uit >40.000 gemeten mRNA's). Van alle geselecteerde mRNA's bleken er geen identieke te zijn tussen alle drie de groepen, en er was slechts een zeer beperkte overlap in twee van de drie groepen. Genexpressie van T-helper cellen lijkt dus niet bruikbaar bij de voorspelling van respons op TNFi therapie.

#### *Met gebruik van eiwitten?*

In hoofdstuk 7 werd gepoogd om met een combinatie van eiwitten en klinische karakteristieken van de patiënt de therapierespons te voorspellen. In een breed spectrum van 57 ontstekings-gerelateerde eiwitten werd in 65 patiënten een model gemaakt dat respons kon voorspellen. In een nieuwe groep van 185 patiënten konden we de resultaten echter niet repliceren. Mogelijk lag dit aan de tijdsduur dat het afgenomen bloed in de vriezer lag opgeslagen voor het werd gemeten, waarin de eiwitpiegels ondanks de lage temperatuur (-80° Celsius) zijn gedaald door afbraak.

### *Met behulp van metaboliëten?*

Metaboliëten, een totaal aan chemische verbindingen in het bloed die verband houdt met allerlei cellulaire processen in het lichaam en voeding, werden in hoofdstuk 8 gemeten om markers te ontdekken die respons konden voorspellen op TNFi. Er werden 139 metaboliëten in het bloed van 124 patiënten gemeten, waaruit voorspellende metaboliëten werden geselecteerd. Vier van deze potentiële metaboliëten konden samen met klinische variabelen een respons op TNFi voorspellen. Deze resultaten werden echter niet gevalideerd in een ander cohort. Daarnaast vonden we veel factoren die van invloed waren op de metaboliëten. Deze factoren bemoeilijkten de interpretatie van de resultaten. Onder andere bleek het gebruik van prednison en soortgelijke therapieën een sterk ontregelend effect te hebben op een groot deel van de onderzochte metaboliëten, vooral op die in de vethuishouding. Deze therapieën en het effect op metaboliëten werden daarom apart onderzocht in hoofdstuk 9. Vrouwelijke RA patiënten toonden de grootste verschillen in metaboliëtenprofiel bij wel of geen gebruik van prednison, wat mogelijk consequenties heeft voor het risico op bijwerkingen zoals hart- en vaatziekten.

### *Met gebruik van DNA?*

In hoofdstuk 10 werd onderzocht of de kans op respons op prednison en methylprednisolon (een korte prednisonkuur per infuus) afhankelijk was van kleine DNA-mutaties in het MDR1-gen. In een kleine groep patiënten leken afwijkingen in dit gen de kans op respons voor methylprednisolon te vergroten. Deze resultaten konden nog niet gerepliceerd worden door het ontbreken van een nieuwe groep patiënten die DNA hebben afgestaan.

## **Discussie: zijn we verdwaald geraakt in replicatie?**

In de afgelopen decennia hebben veel studies naar voorspellers gezocht voor de respons op biologicals in RA. In deze publicaties werd meestal (een set van) voorspellers met potentie voor de klinische praktijk voorgesteld. Deze studies sluiten meestal af met de boodschap dat de enige noodzakelijke stap alvorens volledig gebruik van deze voorspellers, het uitvoeren van een validatie is in een aparte groep patiënten. Dit is gerechtvaardigd, omdat replicatie/validatie een belangrijke procedure is om ervoor te zorgen dat de resultaten reproduceerbaar en biologisch betekenisvol zijn. Helaas worden de meeste resultaten nooit afzonderlijk gevalideerd of genoemd in ander onderzoek. Door dit proces van het systematisch introduceren van niet-geverifieerde resultaten, stapelt de experimentele en ontoepasbare data zich op. Dit kan grote mate van heterogeniteit in de resultaten tussen studies verklaren, die ook gezien werd in het literatuuronderzoek (hoofdstuk 2). Er kan verondersteld worden dat de combinatie van de rapportage- en publicatiebias bijdraagt aan het onderhouden van het proces: elke studie die geen bruikbare voorspellers overhoudt na validatie, moet concurreren met alle (niet-gevalideerde) reeds gepubliceerde positieve resultaten. Het loont daarom om na de ontdekking van voorspellers geen validatie uit te voeren en in het manuscript de nadruk te leggen op deze veelbelovende experimentele resultaten. Met andere woorden, in het academische "publiceer of stap op" -klimaat worden onderzoekers eerder beloond voor de gemakkelijkste en snelste – dan voor de beste weg. Een relevante vraag is dus of de ongevalideerde voorspellers meer zijn dan alleen (publiekelijk gefinancierd) onderzoeksafval. Er bestaat geen twijfel aan de vraag of verbetering in het ontwerp van dit onderzoek nodig is. Het TRIPOD-initiatief heeft als doel om de kwaliteit van prognostische en diagnostische studies wereldwijd te verbeteren (<http://www.tripod-statement.org>). Uiteindelijk is validatie in meerdere cohorten de enige manier om de geloofwaardigheid van een voorspeller aan te tonen, daarbij het credo van prognostisch onderzoek erend: "Replicatie, replicatie en replicatie"

Ondanks de pogingen in dit proefschrift om de resultaten zelf te valideren, konden geen voorspellers gevonden worden die respons voorspellen en gebruikt kunnen worden in de klinische praktijk. Dit betekent niet dat het voorspellen van respons onmogelijk is,

aangezien vele redenen voor het niet kunnen vinden van biologisch relevante voorspellers parallel aan elkaar kunnen bestaan. Het is derhalve denkbaar dat we met een aanpassing van de strategie in dit onderzoeksveld wél reproduceerbare en relevante voorspellers vinden. Belangrijke factoren in deze nieuwe strategie die aandacht verdienen werden onderzocht door interne discussie, deductie en literatuuronderzoek, en worden hieronder nader besproken.

### **Omarm de klinische variatie onder RA-patiënten omdat dit een eerlijke vergelijking van biomarkers mogelijk maakt**

Er bestaat een aanzienlijke variabiliteit in klinische kenmerken tussen RA-patiënten, die wordt beschouwd als één van de hoofdoorzaken van de beperkte overlap in resultaten tussen predictiestudies. In plaats van de variabiliteit in klinische kenmerken te beschouwen als beletsel, kan deze juist worden gebruikt om te corrigeren voor individuele verschillen in biomarkers. Zo kan het bijvoorbeeld voorkomen dat een marker pas bruikbaar is als je het voor mannen en vrouwen apart bekijkt. Een mislukte validatie in een ander cohort kan dus gekoppeld zijn aan een patiëntkarakteristiek die niet gelijk verdeeld is tussen de twee groepen, en pas na correctie voor die factor een valideerbare biomarker oplevert. Door te corrigeren voor een standaard set aan patiëntkarakteristieken, kunnen we wellicht de voorspelling verbeteren en beter vergelijkbaar maken tussen (zeer) verschillende patiëntgroepen.

### **Integratie van systeem-biologie in voorspelmodellen en kwaliteit van de data**

De analyse van grote hoeveelheden data, omvattende alle patiëntkarakteristieken en biomarkers, ligt buiten het bereik van de conventionele medische praktijk. Hoewel in predictiestudies diverse geavanceerde statistische technieken worden toegepast om de meest interessante target(s) uit de data te filteren of om de data te bundelen, is dit in feite meer een simplificatie van de werkelijkheid (data-reductie) dan daadwerkelijke integratie van alle beschikbare informatie. Een systeem-biologische benadering zal daarom meer geschikt zijn om het volledige potentieel van de data te benutten. In de systeembioologie worden alle beschikbare datasets (DNA, RNA, miRNA, proteïne, metabolieten en patiënt karakteristieken) gecombineerd door middel van geavanceerde

wiskundige en computationele technieken, zodat de daadwerkelijk biologische status van de cel bepaald kan worden. Dit veld is echter nog volop in ontwikkeling en systeembiologie wordt nog niet op grote schaal toegepast.

Daarnaast produceren de meettechnieken een grote hoeveelheid data op een relatief kleine hoeveelheid patiëntenmateriaal, waardoor de aanwezigheid van technische variabiliteit of ruis de analyses ernstig verstoort. Ruis kan leiden tot selectie van markers die biologisch minder relevant zijn, ofwel goede markers verhullen. Elk onderzoek moet er daarom naar streven om maximaal te corrigeren voor ruis, door te kiezen voor de beste meettechniek, standaardisering van procedures en zo mogelijk de analyses uit te besteden aan deskundigen, toepassen van geschikte methoden voor de kwaliteitscontrole, en het uitsluiten van monsters met ongeloofwaardig extreme waarden. Ongeacht maximale aanpassingen, zal ruis inherent aan de gebruikte techniek blijven bestaan. Omdat technieken steeds verder ontwikkeld en verbeterd worden, zal de technische variabiliteit uiteindelijk een minder groot probleem worden.

### **Hoge kans op misclassificatie van respons: de noodzaak voor een betere uitkomstmaat**

Hoewel het intuïtief heel duidelijk aanvoelt, een patiënt reageert wél of niet, zit de werkelijke respons er vaak tussenin. Voor een goede definitie moet er vóór en na de behandeling een maat geëvalueerd worden die de ziekte-status van de patiënt uitdrukt. De DAS28 (disease activity score based on 28-joint count) is het meest gebruikte substituut meetinstrument hiervoor. In de DAS28 worden enkele parameters gecombineerd: het aantal gezwollen en/of pijnlijke gewrichten (van 28 vaak aangedane gewrichten), een algehele gezondheidsscore van de patiënt en een ontstekingswaarde uit het bloed. Respons in BiOCURA werd bepaald aan de hand van een combinatie van de verandering in DAS28 over de tijd en de DAS28 op het moment van de respons-meting. Hoewel de DAS28 uitgebreid getest is en een veel gebruikt instrument is in klinische onderzoeken, heeft het zijn nadelen. Een nadeel is, onder andere, dat het een vrij instabiel monitorings-instrument blijkt te zijn, zelfs in RA patiënten die al langere tijd een stabiele ziekte hebben. Dit betekent dat willekeurige schommelingen in ziekteactiviteit optreden, die niet noodzakelijkerwijs een klinische relevantie hebben. Deze schommelingen kunnen derhalve leiden tot een meer willekeurige indeling van responders en niet-responders.

Daarnaast heeft de DAS28 als nadeel dat het de 28 meest aangedane gewrichten meet, en daarmee ook gewrichten overslaat die wél aangedaan kunnen zijn, zoals in de voeten. Patiënten met betrokkenheid van die gewrichten en die klinisch erg verbeteren, worden daarom niet ‘opgepikt’ met de DAS28, en onterecht geclassificeerd als non-responders. Een stabielere en meer objectieve meting van ziekteactiviteit zal in de toekomst ontwikkeld moeten worden om betrouwbaar de respons te kunnen meten.

### **Overweeg andere potentiële voorspellers**

Het is mogelijk dat de biologisch meest betekenisvolle voorspellers voor RA niet in de meest toegankelijke compartimenten, zoals bloed, speeksel en urine, te vinden zijn. Aangezien de ziekte zich vooral manifesteert in de gewrichten, zou het kunnen dat markers gemeten in het gewrichtskapsel of de gewrichtsvloeistof duidelijkere resultaten geven. Vooralsnog hebben studies met metingen van deze materialen echter (ook) geen voorspellers voor respons opgeleverd.

Daarnaast houden predictiestudies vaak geen rekening met het milieu waaraan de patiënt wordt blootgesteld. In BiOCURA werden alleen de toxicologische invloeden van roken en alcohol onderzocht. Er zijn echter veel meer externe factoren die van invloed kunnen zijn, zoals de inhalatie van ‘smog’ en blootstellingen op het werk, socio-economische status, en de darmflora. De invloed van deze factoren op de therapierespons moeten nog onderzocht worden.

### **Predictie kan moeilijker zijn dan vooraf bedacht**

De mogelijkheid bestaat dat voorspellen van de respons te moeilijk is. Niet omdat wij en anderen er niet in geslaagd zijn, maar omdat de biologie van RA mogelijk te complex en variabel is om grip op te krijgen en/of patiënten te onvoorspelbaar reageren op onbekende parameters zoals externe milieufactoren. Zelfs met het ontwikkelen van nieuwe meet- en analysetechnieken en het toenemen van de kennis, staart iedereen zich blind in de chaos die uiteindelijk leidt tot de waarneming van de therapierespons. Er zijn vele parallele omstandigheden in het dagelijks leven, zoals het voorbeeld van Michael Crichton in Jurassic Park (1991), bij het stoten van een bal op een pooltafel en de berekening van waar deze heen gaat. Weliswaar kun je met behulp van de wetten van

Newton waarschijnlijk voorspellen waar deze heen zal gaan, misschien zelfs voor de komende uren. Maar in de praktijk blijkt dat je alleen maar een paar seconden in de toekomst kunt voorspellen, omdat vrijwel onmiddellijk zeer kleine effecten - onvolkomenheden in het oppervlak van de bal of in het hout van de tafel - een verschil beginnen te maken. Zij zullen uiteindelijk zelfs de meest zorgvuldige berekeningen overmeesteren. Een ander voorbeeld, niet losstaand van dit onderzoek, is de rook die stijgt van een sigaret. De rook gaat eerst recht omhoog en spreidt zich dan uit als een diffuse, golvende laag. Maar zelfs de beste supercomputers, die metingen nemen in de meest zorgvuldig gecontroleerde omstandigheden, kunnen niet exact de vorm van de rookwolk voorspellen. Door te experimenteren met de opzet van studies die de werkelijke situatie (sterk) vereenvoudigen, zijn we mogelijk beter in staat om (een deel van de) complexiteit te overwinnen.

### **Hypothesegericht werken**

Binnen RA heerst een algemeen gedachtegoed dat de verandering van RA van een hanteerbaar ziektebeeld naar een ziekte die te genezen is, uiteindelijk pas gerealiseerd kan worden als we alle ziekteprocessen begrijpen. Met voorspelling zou het net zo kunnen zijn. Dit betekent dat energie en geld mogelijk beter besteed kunnen worden aan het begrijpen van de ziekteprocessen in plaats van aan de huidige 'shotgun-benadering', waarin we veel markers tegelijk meten en hopen dat sommigen de roos raken. Met behulp van deze kennis, kunnen hypothesen opgesteld worden die specifiek getest kunnen worden, terwijl ook rekening wordt gehouden met alle (biologisch mogelijke) factoren die van invloed kunnen zijn. Hoewel dit misschien een stap terug lijkt, kan het ons voorwaarts helpen.

### **CONCLUSIE**

Ondanks alle verrichte inspanningen, ontbreken er nog steeds bruikbare voorspellers van therapierespons bij RA. Hoewel de toonaangevende onderzoekers in de reumatologie ervan overtuigd zijn dat voorspellers voor de respons op RA behandeling ooit beschikbaar zullen zijn, zal dit waarschijnlijk een aanpassing in het ontwerp van de predictiestudies vereisen, rekening houdend met de besproken onderwerpen in dit hoofdstuk.

## Dankwoord

Er is een groot aantal personen dat ik graag zou willen bedanken voor hun directe en indirecte bijdrage aan dit proefschrift.

Allereerst mijn promotoren en copromotoren: Floris (Lafeber), je grote wetenschappelijke ervaring in combinatie met je onuitputtelijke creatieve geest waren een verademing om erbij te hebben als het onderzoek vast dreigde te lopen. Daarnaast hielp je pragmatische aanpak als er knopen doorgehakt moesten worden. Ondanks dat RA en statistiek niet je grootste hobby's zijn, boekte het onderzoek zijn vooruitgang en vonden eindproducten hun weg door jouw kwaliteiten. Bedankt voor alle begeleiding. Prof. Bijlsma, beste Hans, u behoort terecht tot één van 's werelds meest gerespecteerde en vooraanstaande wetenschappers en beleidsmakers in de reumatologie. Daarnaast staat u centraal in de samenwerking tussen de reumatologische centra die patiënten hebben geïnccludeerd (de SRU), heeft u alle banden met farmaceuten opgebouwd en warm gehouden en vond u financierders voor mijn aanstelling. Ik ben daarom meer dan vereerd om u als promotor te hebben. Ik hoop dat u komende jaren geniet als president van de Europese reumavereniging, een kroon op uw indrukwekkende carrière. Ruth (Fritsch), je bent een inspiratie als het gaat om het combineren van biochemisch onderzoek, artsenvak en het ouderschap, allemaal uiterst gepassioneerd. Je toewijding aan dit onderzoek was groot, wat bleek als ik weer eens laat in de avond een e-mail met feedback kreeg. Zelfs na je verhuizing naar Oostenrijk bleef je nauw betrokken en enthousiast. Bedankt voor al je input en ik wens je het allerbeste voor jou en je familie. Anne Karien Marijnissen, als opzetter van het onderzoek, binnenhalen van de overheidssubsidie en dagelijks begeleider was je onmisbaar in mijn onderzoek. Je maakte mij vooral wegwijs in de politieke wirwar van partijen, artsen, wetenschappers en journals, met al hun uiteenlopende belangen. Het was erg fijn om te kunnen leunen op je communicatieve vaardigheden.

Ook dank aan alle andere betrokken hoogleraren, assistent professoren en post-docs. Paco, het was een zeer groot genoegen om je aan mijn zijde gehad te hebben, gezien je verregaande epidemiologische kennis en kunde, enorme intelligentie en behulpzaamheid. Ik hoop -en verwacht- dat je het nog ver gaat schoppen in je carrière. Marzia, your molecular biochemical knowledge and experience is larger than anyone else's who helped

me with this research, and I am thankful that I could benefit from it. You are a pure scientist who works extremely hard, best exemplified by your extensive feedback with a total word count larger than the article I send you. Thank you for all your help and I hope your career will flourish. Tim (Radstake), als voorvechter van de systems biology ben je ook één van de meest progressieve wetenschappers die ik ken. Ik hoop dat alle beoogde ontwikkelingen voorspoedig zullen verlopen en je de vraagstellingen uit dit proefschrift in de toekomst kunt beantwoorden. Bedankt voor je input in mijn onderzoek. Jaap (van Laar), evenals ons vorig afdelingshoofd ben je een goede arts, manager én wetenschapper. Je hebt me enorm veel geleerd op het vlak van proper English en publiceren, erg veel dank daarvoor. Hans (Jacobs), een gepassioneerd wetenschapper. Je scherpe blik en ongezouten eerlijkheid maakt elke meeting een stuk waardevoller en productiever, op voorwaarde dat de aanhoorder op de juiste golflengte afstemt. Toen ik dat eenmaal door had bleek je een grote hulp. Bedankt voor alle input. René (Eijkemans), bij je oratie kwam ik erachter dat jouw visie op biostatistiek was wat ik al langer ervaarde. Je gaf me de insteek en woorden voor het schrijven van mijn thesis, en per slot van rekening werd je ook nog voorzitter van de lezerscommissie. Ik hoop dat je nog veel anderen met je visie aansteekt. Team Analytical Biosciences from Leiden, including Herman, Amy, Thomas, Ruud, Jan and Slavik, thank you for all the help in dealing with the complexity of metabolites. It was a pleasure to experience your complete determination and fascination with the metabolic state of human beings. Junzeng, you have a great working spirit and know a lot about metabolites, I could not wish for a better partner to write the two difficult articles with. Victor Jong, thank you for showing me how big data can both be easy and difficult. Erik (Hack), Wilco en de rest van de Luminex core facility, samen zijn we er nogmaals achter gekomen hoe ingewikkeld en beïnvloedbaar biomarkermetingen kunnen zijn. Bedankt voor het blijven samenwerken en uiteindelijk helpen aan betrouwbare data. Maarten Kraan, Katerina, Marita and Linda, thank you for the opportunity to also study DNA by measuring samples at your labs at AstraZeneca in Sweden.

Reumatologen van het UMCU, bedankt voor het opzetten van de BiOCURA studie en includeren van de enorme aantallen patiënten. Het cliché is waar, zonder jullie was dit

onderzoek nooit mogelijk geweest. Dank aan Evelien, Aike, Jocea, Hans B, Hans J, Helen, Ruth, en alles AIOS. In het bijzonder, Janneke, bedankt voor het helpen opzetten van de BiOCURA studie ver voor mijn tijd. Ook alle reumatologen van de Stichting Reuma-onderzoek Utrecht (SRU), eveneens dank voor alle patiënten die jullie hebben geïnccludeerd tijdens de volle spreekuren. Jullie hulp is voor alle SRU-studies van ongekend belang. Bij naam wil ik de contactpersonen van deze ziekenhuizen noemen en zij die meegeschreven hebben aan artikelen: Martijn Custers, Willemijn Noort-van der Laan, Evert-Jan ter Borg, Judith Peeters, Suzanne Linn-Rasker, Yolande Schenk, Dudy Kuiper-Geertsma, Marita Nabibux en Karin Verhoef.

Alle research verpleegkundigen, Karin, Annemiek, Joke en Marieke, de bijdrage die jullie geleverd hebben aan dit- en alle andere onderzoeken van de afdeling - is indrukwekkend. Bedankt voor de lieve behandeling van patiënten en alle gewrichten die jullie geteld hebben: in mijn snelle berekening zeker meer dan 50.000. Wat is de afdeling (terecht) trots op jullie om te hebben. Marieke, extra dank voor je zorgvuldige data-invoer en blik op de lay-out van dit proefschrift.

Alle laboranten, Arno, Katja, Kim, Dorien en Marion, jullie hebben enorm veel samples verwerkt en opgeslagen, alles zorgvuldig geprotocolleerd. De hoeveelheid plastic buisjes met uitgesplitst serum, RNA, cellen en urine voor deze studie is duizelingwekkend en overstijgt met gemak de 3 miljoen. Dat kan alleen met een enorm sterk team, vele uren werk en een strenge baas.

Alle andere PhD'ers en overige collega's, dank voor de goede werksfeer, de gedeelde passie voor goede wetenschap met oneindig veel (positieve) feedback op elkaars werk en de minder serieuze momenten. Xavier, fellow-epi, sterkte met alle frustraties die altijd op jouw pad komen: krappe bureaus, viseters, bananendieven en mensen die je analyses anders willen. Maud, sinds je op mijn eerste werkdag extra vroeg was gekomen om mijn bureau schoon te maken wist ik dat ik veel aan je zou hebben. Dank voor de uitgebreide introductie in de wondere wereld van promoveren. Sandhya, thank you for teaching me to speak proper English. Jan-Ton, deeltijdpromovendus, op dat je een top orthopeed mag

worden. Thijmen, ik snap nog steeds niet wat een ingenieur in het ziekenhuis moet, maar koffie zetten kun je als de beste. Nick, ooit ga ik je woordgrappen missen. Nee echt. Huub, volgens mij ben ik even groot als jij. Astrid, je idee voor een luisterboek-versie van dit proefschrift was briljant. Jelena, you're also fun when you're hangry. Eefje, maat, fijn dat je op kwam voor Nijmegen. Simon, als je pensionneert, begin dan je eigen pubquiz. Je weet te veel. Erwin, ben je nog steeds fan van de traditionele Franse keuken? Lize, ik ben nog altijd onder de indruk van je elektrisch bestuurbare gordijnen. Anne, ik heb veel met je gelachen op onze tripjes, goede herinneringen aan. Marjolein, Harrie, en Nicolien, dank voor jullie grote input bij de research meetings. Alle andere PhD-students, Mary, Willem-Paul, Bernadette en Roos, ik hoop dat jullie mooie proefschriften tot stand gaan brengen. Ook mijn twee masterstudenten, Jan en Bery, bedankt voor al jullie werk. Researchers of Tims group, Elena, Sofie, Luuk, Maarten, Frederique, and Sandra, thank you for sharing your thoughts on miRNA and mRNA experiments and analyses. In special, Cornelis, you have been a great mentor in the lab explaining me and Arno all about the practical molecular work. Joel and Aridaman, thank you for the fruitful discussions on my data, it really helped me to gain insights from other points of view. Zoveel meer (oud)collega's van de afdeling reumatologie heb ik nog niet genoemd: Ron (Derksen), je colleges tijdens de geneeskundestudie waren legendarisch en als stagebegeleider tijdens mijn onderzoek naar lupus heb je me weten te enthousiasmeren voor de immunologie. Remco, als stagebegeleider en zaalsupervisor bij de semi-artsstage heb je mij geïnspireerd. Lucy en Judy, het was een genoegen om de sponsoring en teams voor de singelloop te organiseren. Monique en Anja, bedankt voor het afhandelen van de stroom aan verzoeken voor onmogelijke vergaderingen, die wonderbaarlijk genoeg toch vaak hebben plaatsgevonden. Alle andere VRCs, managers, etc. die ik nog niet gehad heb, eveneens bedankt voor de leuke tijd.

Ik wil ook mijn familie en vrienden bedanken voor de support. Allereerst mijn ouders voor een bijzonder plezierige en veilige jeugd, de geboden mogelijkheden en het mateloze vertrouwen in hun kinderen. Mijn zussen, broer en hun aanhang, Roos & Arjen, Laura & Hizkia, Eva en Coen, waardoor het thuis altijd bruist. Opa, dank voor de genen die ons nieuwsgierig maken. Tamarah, bedankt voor alle leuke momenten en steun in het laatste

jaar. Daarnaast ook dank aan enkele vrienden. Jeroen van Duijn, voor alle jaren vóór en nadat ons huis afbrandde. Babette, voor je ongeremde vrolijkheid, waar je ook verschijnt. Mignon, Sophie, Martine en Bastiaan, voor de vakanties, etentjes, interviews, en eveneens Bastiaan omdat je mij enthousiast hebt weten te maken voor de reumatologie. Maurice, Rasmus, Jeroen, Teuni, Thomas, Gerrit en Nina, dat we nog lang met elkaar mogen optrekken. Anjes, Yoram, Jeffrey en Daan, voor het samen spelen, onze muziekproducties en de optredens. Rutger, voor het oprichten van onze eervolle VOF. Rick, voor je prachtige ontwerp van de cover.

Mijn paranimfen, Arno en Maurice, bedankt voor de support tijdens het onderzoek en bij de verdediging. Arno, jij hebt bijzonder veel werk verzet voor deze studie, met name bij het opzetten en uitvoeren van alle protocollen omtrent de sample collectie en verwerking. Je noemt de BiOCURA-studie vaak liefkozend je 'kindje'. Ik ben blij dat ons gedeeld ouderschap al die jaren standgehouden heeft. Maurice, we hebben samen kennis gemaakt met de wondere wereld van onderzoek toen we werkten aan ons profielwerkstuk. Leuk dat de wetenschap ons heeft weten te boeien tot zoveel jaren daarna, en we hier nog steeds uren over kunnen praten.

Graag bedankt ik ook alle sponsors van dit onderzoek, ZonMW, Vrienden van UMC Utrecht en in het bijzonder Jo en Harrie: tijdens het onderzoek heb ik mij vaak 'jullie onderzoeker' gevoeld. Zonder jullie steun had ik niet kunnen werken aan dit onderzoek, en daar ben ik erg dankbaar voor. Ik heb genoten van onze bijeenkomsten in het ziekenhuis en bij jullie in het prachtige Brabant, want uiteindelijk gaat dit onderzoek over gezond en vitaal je eigen leven kunnen leiden. Dat doen jullie gelukkig nog altijd.

Dan rest mij alleen nog alle patiënten te bedanken die belangeloos hebben bijgedragen aan dit onderzoek. Uw enorme bijdrage qua tijd en moeite, maar ook uw overgave en het vrijwillig opgeven van privacy, zijn bewonderingswaardig. Ik neem er mijn pet diep voor af, bedankt. Ik hoop dat u allen, uiteindelijk, een goede responder zult zijn.

## **Curriculum vitae**

Bart was born on June 17th 1988 in Nijmegen. Eighteen years later in June 2006, he graduated *cum laude* from secondary school, the SSgN in his hometown. Bart studied medicine at the Utrecht University, focusing his final year mainly on rheumatology and clinical immunology. After his graduation in April 2013 and intrigued by the complexity and elusiveness of rheumatic diseases, Bart began a PhD trajectory at the Rheumatology and Clinical Immunology department of the University Medical Center Utrecht. In the following years, as additional methodological capacities were required to maximize the scientific potential of his research, Bart earned a master's degree epidemiology at the Vrije University in Amsterdam in December 2015. Finally, his thesis on prediction of therapeutic response in rheumatoid arthritis was defended on March 23<sup>rd</sup>, 2017.

## List of publications

**Cuppen BV**, Pardali K, Kraan MC, Marijnissen AC, Yrlid L, Olsson M, Bijlsma JW, Lafeber FP, Fritsch-Stork RD. Polymorphisms in the multidrug-resistance 1 gene related to glucocorticoid response in rheumatoid arthritis treatment. *Rheumatol Int.* 2017 Jan 28. [Epub ahead of print]

**Cuppen BV**, Jacobs JW, Ter Borg EJ, Marijnissen AC, Bijlsma JW, Lafeber FP, van Laar JM; all SRU investigators. Necessity of TNF-alpha inhibitor discontinuation in rheumatoid arthritis is predicted by smoking and number of previously used biological DMARDs. *Clin Exp Rheumatol.* 2016 Oct 7. [Epub ahead of print]

Fu J, **Cuppen BV**, Welsing PM, van Wietmarschen H, Harms AC, Berger R, Koval S, Fritsch-Stork RD, Bijlsma JW, Hankemeier T, van der Greef J, Lafeber FP. Differences between serum polar lipid profiles of male and female rheumatoid arthritis patients in response to glucocorticoid treatment. *Inflammopharmacology.* 2016 Dec;24(6):397-402.

**Cuppen BV**, Fu J, van Wietmarschen HA, Harms AC, Koval S, Marijnissen AC, Peeters JJ, Bijlsma JW, Tekstra J, van Laar JM, Hankemeier T, Lafeber FP, van der Greef J; all Society for Rheumatology Research Utrecht investigators. Exploring the Inflammatory Metabolomic Profile to Predict Response to TNF- $\alpha$  Inhibitors in Rheumatoid Arthritis. *PLoS One.* 2016 Sep 15;11(9):e0163087.

Fu J, Schoeman JC, Harms AC, van Wietmarschen HA, Vreeken RJ, Berger R, **Cuppen BV**, Lafeber FP, van der Greef J, Hankemeier T. Metabolomics profiling of the free and total oxidised lipids in urine by LC-MS/MS: application in patients with rheumatoid arthritis. *Anal Bioanal Chem.* 2016 Sep;408(23):6307-19.

**Cuppen BV**, Rossato M, Fritsch-Stork R, Concepcion AN, Schenk Y, Bijlsma JWJ, Radstake TR, Lafeber P, all Society for Rheumatology Research Utrecht investigators. Can baseline serum microRNAs predict response to TNF-alpha inhibitors in rheumatoid arthritis?. *Arthritis Res Ther.* 2016 Aug 24;18:189

**Cuppen BV**, Welsing PM, Sprengers JJ, Bijlsma JW, Marijnissen AC, Laar JM, Lafeber FP, Nair SC. Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability. *Rheumatology (Oxford)*. 2016 May;55(5):826-39.

Jurgens MS, Overman CL, Jacobs JW, Geenen R, **Cuppen BV**, Marijnissen AC, Bijlsma JW, Welsing PM, Lafeber FP, van Laar JM. Contribution of the subjective components of the Disease Activity Score to the response to Biological Treatment in Rheumatoid Arthritis. *Arthritis Care Res (Hoboken)*. 2015 Jul;67(7):923-8.

Luijten RK, **Cuppen BV**, Bijlsma JW, Derksen RH. Serious infections in systemic lupus erythematosus with a focus on pneumococcal infections. *Lupus*. 2014 Dec;23(14):1512-6.

### **Conference abstracts**

A proteomics approach to predict the TNF-alpha inhibitor response in RA: the added clinical value of a protein score. European League Against Rheumatism (EULAR) congress 2015. (Oral presentation)

Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability. European League Against Rheumatism (EULAR) congress 2015. (Poster presentation)

Towards individualized risk determination in RA: A prediction model for TNFi discontinuation within the first year after start. European League Against Rheumatism (EULAR) congress 2015. (Poster presentation)

Predictie van respons op biologicals in Reumatoïde Artritis. ZonMw Goed Gebruik Geneesmiddelen congres 2015. (Oral presentation)

Proteomics predict treatment response in RA for both EULAR and newly proposed response criteria. Nederlandse Vereniging voor Reumatologie (NVR) najaarsdagen 2014. (Oral presentation)

Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability. Nederlandse Vereniging voor Reumatologie (NVR) najaarsdagen 2014. (Poster presentation)