Predictive value of serum calprotectin (S100A8/A9) for clinical response after starting or tapering anti-TNF treatment in patients with rheumatoid arthritis

Lieke Tweehuysen,¹ Nathan den Broeder,¹ Noortje van Herwaarden,¹ Leo A B Joosten,² Peter L van Lent,³ Thomas Vogl,⁴ Frank H J van den Hoogen,¹,⁵ Rogier M Thurlings,⁵ Alfons A den Broeder¹,⁵

ABSTRACT

Objectives Calprotectin (S100A8/A9) has been correlated with disease activity in rheumatoid arthritis (RA). The aim of this study was to investigate the predictive value of serum calprotectin for clinical response after starting and tapering anti-tumour necrosis factor treatment in RA.

Methods Serum samples and clinical outcomes were derived from two longitudinal RA studies. At baseline (starting or tapering of adalimumab or etanercept), calprotectin levels were determined by ELISA. In the Biologic Individual Optimised Treatment Outcome Prediction (BIO-TOP) study, treatment effect was assessed after 6 months using the European League Against Rheumatism (EULAR) response criteria. In the Dose Reduction Strategies of Subcutaneous TNF Inhibitors (DRESS) study, patients were classified at 18 months as being successfully dose reduced, discontinued or not able to reduce the dose. Area under the receiver operating characteristic curves (AUC) were generated to evaluate the predictive value of calprotectin and logistic prediction models were created to assess its added value.

Results In the BIO-TOP study, calprotectin levels were higher in responders (n=50: 985 ng/mL (p25–p75: 558–1417)) compared with non-responders (n=75: 645 ng/mL (p25–p75: 415–973), p=0.04). AUC for predicting EULAR good response was 0.61 (95% CI 0.50 to 0.71). The prediction model with calprotectin (AUC 0.77, 95% CI 0.68 to 0.85) performed similarly to the baseline model (AUC 0.74, 95% CI 0.65 to 0.82, p=0.29). In the DRESS study, calprotectin levels were similar between the three groups (n=47; n=19; n=36) and calprotectin was not predictive for clinical response after tapering.

Conclusions Serum calprotectin has some predictive value for clinical response after starting anti-TNF treatment, although it has no added value to other clinical factors. In patients with low disease activity, serum calprotectin is not predictive for clinical response after tapering anti-TNF treatment.

Trial registration number NTR4647 (BIO-TOP study) and NTR3216 (DRESS study); Pre-results.

INTRODUCTION

The introduction of biological disease-modifying antirheumatic drugs (bDMARD) has improved the treatment outcomes of rheumatoid arthritis (RA). However, approximately 60% of patients with RA do not achieve good clinical response after 6 months of treatment with a bDMARD, including tumour necrosis factor inhibitors (TNFi). Further, tapering TNFi has been shown feasible in a large proportion of patients with RA with low disease activity but again, not in all patients. Non-responding after starting a TNFi or flaring after tapering of a TNFi are both undesirable, since a (short) period of high disease activity might cause worsening of physical functioning and radiographic joint progression.

The ability to accurately predict individual response after starting or tapering of a TNFi might improve treatment outcomes compared with the current ‘trial-and-error’ approach.
treatment. In patients who are unlikely to respond to (a certain) TNFi, starting another TNFi or bDMARD with another mode of action might potentially be more effective. In addition, when it can be predicted that tapering will be unsuccessful in a patient, tapering would not be attempted thereby preventing disease flares, minimising physician efforts and easing uncertainty in patients. Notably, in both scenarios, the gains would be better treatment outcomes, not lower direct costs per se.

Calprotectin (also known as S100A8/A9 and MRP8/14) might be a promising biomarker to predict clinical response to anti-TNF treatment. In contrast to acute phase proteins which are mainly of hepatic origin, calprotectin is released predominantly by granulocytes at sites of inflammation.5 It also diffuses easily from inflamed joints into the blood circulation because of its relatively low molecular weight.6

Previous studies have shown that serum calprotectin is cross-sectionally correlated with clinical and laboratory markers of disease activity in RA.7–11 Studies that have investigated the longitudinal predictive value of calprotectin for clinical response to bDMARD treatment show however conflicting results. Choi et al demonstrated that serum calprotectin at baseline predicts response to treatment with respectively adalimumab (ADA), infliximab and rituximab.12 However, in other studies baseline calprotectin could not predict responsiveness to treatment with a bDMARD.8–10 On the other hand, decreased calprotectin levels after 4 weeks of bDMARD treatment were consistently predictive of clinical response.12–15 And in patients with juvenile idiopathic arthritis (JIA) in clinical remission, high calprotectin levels at the moment of discontinuation of etanercept (ETN) were associated with subsequent flare.16 This association has not yet been investigated in RA.

Before serum calprotectin can be used to guide clinical decision-making in daily practice, it must meet minimal validation criteria, as proposed by the OMERACT Soluble Biomarker Group.17 Furthermore, the prediction of the clinical outcome should have therapeutic consequences. Achieving good response after starting treatment and maintaining response after tapering treatment are both clinical important scenarios with uncertainty on clinical outcome. Therefore, the aim of this study was to investigate the predictive value of serum calprotectin for clinical response after starting and tapering anti-TNF treatment in patients with RA.

METHODS
Study population
Baseline serum samples and clinical outcomes of patients with RA starting ADA or ETN were derived from the Biologic Individual Optimised Treatment Outcome Prediction (BIO-TOP) study (Dutch trial register, NTR5216).18 In this prospective longitudinal prediction study, patients with RA > 18 years starting with or switching to a bDMARD were enrolled between 2014 and 2016. The same data were collected of patients with RA included in the dose tapering arm of the Dose Reduction Strategies of Subcutaneous TNF Inhibitors (DRESS) study (Dutch trial register, NTR3216)19, an 18-month open randomised clinical trial investigating non-inferiority of a dose reduction strategy of ADA or ETN compared with usual care. In the DRESS study, patients with RA using ADA or ETN at any stable dose and interval for at least 6 months with stable low disease activity at two subsequent visits were enrolled in 2011 and 2012. Full details of this study have been reported previously.20 These two studies were performed in two hospitals in the Netherlands (Sint Maartenskliniek Nijmegen and Maartenskliniek Woerden) and were both approved by the local ethics committee (CRO region Arnhem-Nijmegen, NL47946.091.14 and NL37704.091.11). Procedures followed were in accordance with the Declaration of Helsinki and all patients gave written informed consent.

Clinical assessments
In the BIO-TOP study, treatment effect was assessed at month 6 with the commonly used European League Against Rheumatism (EULAR) response criteria (good vs moderate/no response).21 In the DRESS study, three clinical outcomes were defined at month 18: successful dose reduction, successful discontinuation and not able to reduce dose. Successful dose reduction was defined as using the TNFi at a longer interval or lower dose than at enrolment with concurrent low disease activity. In both studies, calprotectin levels were not available when disease activity was assessed, preventing expectation bias.

Serum calprotectin measurement
At baseline (starting or tapering of ADA or ETN) serum samples were collected and centrifuged within 1 hour. Subsequently the samples were stored at −80°C until analysis took place (October 2016). Calprotectin levels were measured in the laboratory of the University of Münster using an ELISA, as described previously.22 The readers of the assay were blinded for the disease activity and medication use of patients.

Statistical analysis
Descriptive statistics are reported as either mean (±SD), median (IQR (p25–p75)) or frequency depending on data distribution. Baseline characteristics were compared using Student’s t-test (or, if not normally distributed, Wilcoxon rank-sum test) and χ2 test for continuous and categorical data, respectively. Correlations between calprotectin and clinical variables (ie, age, gender, disease duration, rheumatoid factor (RF), anti-citrullinated protein antibodies, 28-joint count Disease Activity Score using C-reactive protein (DAS28-CRP) and its components, CRP and erythrocyte sedimentation rate (ESR)) were cross-sectionally explored by Spearman correlation analysis. In the BIO-TOP study, discontinuation of the TNFi before 6 months due to lack of effect was regarded as non-response and for discontinuation...
due to other reasons the clinical response at month 3 was carried forward.

Area under the receiver operating characteristic curves (AUC) were generated to evaluate the predictive value of baseline serum calprotectin levels for respectively EULAR good response (yes vs no), successful dose reduction (yes vs no), successful discontinuation (yes vs no) and not able to reduce dose (yes vs no). For AUCs with a 95% CI lower bound ≥0.5, we additionally performed logistic prediction modelling. First, univariate logistic regression analyses were performed to assess which demographic, disease-specific and treatment-specific variables at baseline were associated with the treatment outcome. Variables that showed an association in the univariate analyses were entered in a multivariate logistic regression analysis using stepwise backwards selection to construct a baseline prediction model taking into account the rule of thumb of 1 predictor per 10 patients. Subsequently, we added calprotectin to the baseline model and tested the equality of the two AUCs by using the algorithm suggested by DeLong et al.23 Missing data in this analysis were addressed using multiple imputation. A p value <0.05 was considered statistically significant. All analyses were performed using STATA V.13.1.

RESULTS

Patient characteristics

Baseline serum samples and 6-month clinical outcome data were available of 50 patients starting ADA and 75 patients starting ETN in the BIO-TOP study. Additionally, 102 of 121 (84%) patients randomised to the tapering arm of the DRESS study had available baseline serum samples of whom 38 patients were treated with ADA and 64 patients with ETN. Baseline characteristics of the included patients in both cohorts are summarised in Table 1.

Serum calprotectin at baseline was similar between patients who started TNFi and patients who tapered TNFi: 680 ng/mL (p25–p75: 433–1252) vs 612 ng/mL (p25–p75: 475–927) (p=0.15). In the BIO-TOP study, calprotectin levels were weakly to moderately significantly correlated with DAS28-CRP (r=0.32, p<0.001), ESR (r=0.41, p<0.001) and CRP (r=0.57, p<0.001) at baseline. Also, calprotectin was significantly higher in RF-positive patients (r=0.19, p=0.03). In the DRESS study, calprotectin levels were only significantly correlated with CRP (r=0.21, p=0.03).

Baseline calprotectin levels and correlation with response to treatment

Fifty of 125 (40%) patients starting a TNFi achieved EULAR good response at month 6. Baseline calprotectin levels were higher in responders (985 ng/mL (p25–p75: 558–1417)) compared with non-responders (645 ng/mL (p25–p75: 415–973), p=0.04) (figure 1, left panel). Responders also had a higher DAS28-CRP (4.4 (SD 0.8) vs 3.8 (SD 1.2), p=0.003) and a higher CRP (8 mg/L (p25–p75: 2–25) vs 3 mg/L (p25–p75: 0–14), p=0.008) at baseline. More patients achieved good response after ETN treatment (38 of 75 (51%) patients) compared with ADA treatment (12 of 50 (24%) patients) (p=0.003). This

Table 1 Baseline characteristics

<table>
<thead>
<tr>
<th>Table 1 Baseline characteristics</th>
<th>BIO-TOP study n=125</th>
<th>DRESS study n=102</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)*</td>
<td>57 (12)</td>
<td>59 (10)</td>
</tr>
<tr>
<td>Female gender</td>
<td>81 (65)</td>
<td>62 (61)</td>
</tr>
<tr>
<td>Disease duration (years)†</td>
<td>4 (1–10)</td>
<td>11 (6–17)</td>
</tr>
<tr>
<td>RF positive</td>
<td>74/123 (60)</td>
<td>80 (78)</td>
</tr>
<tr>
<td>ACPA positive</td>
<td>65/116 (56)</td>
<td>73 (72)</td>
</tr>
<tr>
<td>Disease characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS28-CRP*</td>
<td>4.0 (1.1)</td>
<td>2.2 (0.6)</td>
</tr>
<tr>
<td>TJC†</td>
<td>4 (2–9)</td>
<td>0 (0–1)</td>
</tr>
<tr>
<td>SJC†</td>
<td>4 (1–7)</td>
<td>0 (0–0)</td>
</tr>
<tr>
<td>PGA, VAS 0–100 mm†</td>
<td>62 (20)</td>
<td>23 (17)</td>
</tr>
<tr>
<td>CRP (mg/L)†</td>
<td>5 (1–19)</td>
<td>3 (3)</td>
</tr>
<tr>
<td>ESR (mm/hour)†</td>
<td>17 (7–31)</td>
<td>12 (7–20)</td>
</tr>
<tr>
<td>Calprotectin (ng/mL)†</td>
<td>680 (433–1252)</td>
<td>612 (475–927)</td>
</tr>
<tr>
<td>Treatment characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of previous bDMARDs†</td>
<td>0 (0–1)</td>
<td>0 (0–1)</td>
</tr>
<tr>
<td>Current TNFi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADA</td>
<td>50 (40)</td>
<td>38 (37)</td>
</tr>
<tr>
<td>ETN</td>
<td>75 (60)</td>
<td>64 (63)</td>
</tr>
<tr>
<td>Duration current TNFi (years)†</td>
<td>NA</td>
<td>3 (2–6)</td>
</tr>
<tr>
<td>Concomitant treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csDMARDs</td>
<td>97 (78)</td>
<td>59 (58)</td>
</tr>
<tr>
<td>MTX</td>
<td>66 (53)</td>
<td>46 (45)</td>
</tr>
<tr>
<td>NSAIDs</td>
<td>80 (64)</td>
<td>57 (56)</td>
</tr>
<tr>
<td>Oral glucocorticoids</td>
<td>22 (18)</td>
<td>5 (5)</td>
</tr>
</tbody>
</table>

Data presented as number (%) unless otherwise noted. *Mean (SD). †Median (p25–p75). ‡Missing data BIO-TOP study; in eight patients (6%) PGA is missing and in nine patients (7%) ESR is missing. If PGA was missing, DAS28-CRP was calculated with three variables: TJC, SJC and CRP.

ACPA, anti-citrullinated protein antibodies; ADA, adalimumab; bDMARD, biological disease-modifying antirheumatic drug; BIO-TOP, Biologic Individual Optimised Treatment Outcome Prediction; CRP, C-reactive protein; csDMARD, conventional synthetic DMARD; DAS28-CRP, 28-joint count Disease Activity Score using CRP; DRESS, Dose Reduction Strategies of Subcutaneous TNF Inhibitors; ESR, erythrocyte sedimentation rate; ETN, etanercept; MTX, methotrexate; NA, not applicable; NSAID, non-steroidal anti-inflammatory drug; PGA, patient global assessment of disease activity; RF, rheumatoid factor; SJC, swollen joint count; TJC, tender joint count; TNFi, tumour necrosis factor inhibitor; VAS, visual analogue scale.

*Mean (SD). †Median (p25–p75). ‡Missing data BIO-TOP study; in eight patients (6%) PGA is missing and in nine patients (7%) ESR is missing. If PGA was missing, DAS28-CRP was calculated with three variables: TJC, SJC and CRP.

ACPA, anti-citrullinated protein antibodies; ADA, adalimumab; bDMARD, biological disease-modifying antirheumatic drug; BIO-TOP, Biologic Individual Optimised Treatment Outcome Prediction; CRP, C-reactive protein; csDMARD, conventional synthetic DMARD; DAS28-CRP, 28-joint count Disease Activity Score using CRP; DRESS, Dose Reduction Strategies of Subcutaneous TNF Inhibitors; ESR, erythrocyte sedimentation rate; ETN, etanercept; MTX, methotrexate; NA, not applicable; NSAID, non-steroidal anti-inflammatory drug; PGA, patient global assessment of disease activity; RF, rheumatoid factor; SJC, swollen joint count; TJC, tender joint count; TNFi, tumour necrosis factor inhibitor; VAS, visual analogue scale.

can be explained by the fact that in our hospital ADA was reserved for patients who had failed on ETN treatment. As a consequence, patients treated with ADA were presumably more refractory to treatment with a TNFi.

Of the patients who tapered ADA or ETN, 47 (46%) successfully reduced the TNFi dose, 19 (19%) successfully discontinued the TNFi and 36 (35%) could not reduce the TNFi dose. Calprotectin levels at baseline were similar between these groups: 599 ng/mL (p25–p75: 473–965), 629 ng/mL (p25–p75: 454–896) and 624 ng/mL (p25–p75: 514–931) (p=0.80) (figure 1, right panel). The patient group that could not reduce the dose had a lower percentage of RF-positive patients (24 of 36 (67%) vs 56 of 66 (85%), p=0.03) and had a near significantly higher DAS28-CRP at baseline (2.4 (SD 0.7) vs 2.1 (SD 0.6), difference +0.3 (95% CI −0.001 to 0.51)).

Predictive value of calprotectin for clinical response after starting a TNFi

The AUC for predicting EULAR good response versus EULAR moderate/no response using baseline serum calprotectin was 0.61 (95% CI 0.50 to 0.71), while for respectively baseline DAS28-CRP and CRP AUCs were: 0.68 (95% CI 0.58 to 0.77) and 0.64 (95% CI 0.54 to 0.74).

Subgroup analyses were performed for serum calprotectin, in which the AUC tended to be greater in patients who started treatment with ADA (0.68, 95% CI 0.49 to 0.88) compared with patients who started ETN (0.49, 95% CI 0.36 to 0.63) (p=0.11). There were no statistically significant differences in AUCs for concomitant conventional synthetic DMARD (csDMARD) use (yes vs no) or oral glucocorticoid use (yes vs no).

The prediction model with calprotectin (AUC 0.77, 95% CI 0.68 to 0.85) performed similarly to the baseline model (backward selected variables: patient global assessment of disease activity, DAS28-CRP, used TNFi (ADA vs ETN) and the interaction between calprotectin and used TNFi) (AUC 0.74, 95% CI 0.65 to 0.82, p=0.29) (Figure 2).

Predictive value of calprotectin for clinical response after tapering a TNFi

Serum calprotectin was not predictive for clinical response after tapering anti-TNF treatment. AUCs were 0.52 (95% CI 0.40 to 0.63) for predicting successful dose reduction, 0.53 (95% CI 0.39 to 0.67) for successful discontinuation and 0.54 (95% CI 0.42 to 0.66) for not able to reduce dose. Also, DAS28-CRP and CRP were not predictive for the three treatment outcomes with corresponding AUCs varying from 0.54 to 0.61 for DAS28-CRP and from 0.46 to 0.56 for CRP. Subgroup analyses showed no statistically significant differences in AUCs for used TNFi (ADA vs ETN), concomitant csDMARD use (yes vs no) or oral glucocorticoid use (yes vs no). Logistic prediction modelling was deemed unnecessary, since the univariate AUCs already demonstrated no predictive value.

DISCUSSION

In this study, we have investigated the predictive value of baseline serum calprotectin for clinical response after both starting and tapering of anti-TNF treatment in patients with RA in daily practice. We have shown that serum calprotectin has some predictive value for clinical response after starting anti-TNF treatment, although it has no added value to other clinical factors, such as DAS28-CRP. Moreover, serum calprotectin was not predictive for clinical response after tapering anti-TNF
treatment in patients with low disease activity. We would like to discuss our findings in more detail below.

Our study supports the previous findings of a meta-analysis which showed that serum calprotectin levels were positively correlated to clinical disease activity measures including DAS28-CRP and CRP. However, in our study calprotectin levels at baseline were not significantly higher in the patient group with high disease activity that started a TNFi compared with the patient group with low disease activity that tapered a TNFi. Second, we could confirm in accordance with a previous study that baseline serum calprotectin levels were higher in responders after starting a TNFi compared with non-responders. Next, we demonstrated that calprotectin has some predictive value for clinical response after starting a TNFi, but has no added value to clinical factors routinely measured in RA in daily practice, such as DAS28-CRP. This lack of added value of calprotectin was observed before in prediction of response after starting a TNFi in JIA. Thus, measuring baseline serum calprotectin levels seems unable to help optimising therapeutic decision-making in daily practice due to its substantial overlap with already used disease activity measures (DAS28-CRP and CRP).

Calprotectin tended to be a better predictor for response to ADA treatment than ETN treatment, but this was not significant due to the lower group sizes. Since ETN was the first-choice bDMARD at our hospital, this finding might indicate that a serum calprotectin level without previous treatment with a TNFi has less predictive value than a serum calprotectin level after previous treatment with a TNFi for clinical response to the next TNFi. This seems to match with previous findings in other studies showing that decreased calprotectin levels after 4 weeks of bDMARD treatment were consistently predictive of clinical response to the same bDMARD.

In patients who tapered treatment with a TNFi, we found no differences in baseline calprotectin levels between the three treatment outcome groups. Also, calprotectin had similar AUC characteristics as DAS28-CRP and CRP. We assume that calprotectin behaves like an acute phase protein. Since tapering was only performed in patients with a low disease activity, there was correspondingly a low intervariability in CRP and calprotectin levels at baseline. As a result, measuring serum calprotectin levels seems unable to distinguish patients who are doing well while using a TNFi from those who are doing well because of a TNFi.

A strength of our study is that we investigated the predictive value of serum calprotectin for clinical response after both starting and tapering of ADA and ETN in a large number of patients with RA treated in the same hospital. We specifically focused on the added value of calprotectin for starting a TNFi and to the best of our knowledge were the first to investigate serum calprotectin for clinical response after tapering a TNFi in RA.

A limitation of our study might be that we have not longitudinally measured calprotectin levels after starting or tapering TNFi. Although other studies have shown that decreased calprotectin levels after 4 weeks of bDMARD treatment were predictive of clinical response, we believe that a biomarker measured after initiating starting or tapering of a bDMARD is less relevant, because the treatment decision is then already made and future treatment decisions can be based on actual clinical response rather than changes in biomarkers. Preferably, a biomarker would be identified that can prevent starting
or tapering of a bDMARD in patients who are likely to fail on it.

Another limitation might be the considerable coefficient of variation (CV) for calprotectin due to its limited linearity range.\(^7\) We tried to minimise it by measuring each sample in three different dilutions and accepting results as reliable if all values coincided after recalculation. There is conflicting evidence whether serum or plasma is preferable for the measurement of calprotectin. A systematic review has demonstrated that the CV for calprotectin is lowest in studies using serum, but Nordal et al found similar variability in calprotectin levels when measured in plasma and serum.\(^{25}\)

Also, calprotectin levels are not a specific marker for RA activity. Calprotectin levels might be affected by the presence of cardiovascular disease and obesity complicating prediction of clinical response to treatment of RA.\(^{26}\)\(^{27}\)

Furthermore, we included only patients with established RA and therefore our results might not be valid for patients with early RA, especially since it has been shown that calprotectin is a better predictor of response to methotrexate therapy in patients with early RA (<1 year) as compared with later onset of disease (>1 year) (10). This difference might be explained by the fact that neutrophil granulocytes are more prevalent in early RA.\(^{28}\) However, the heterogeneity in treatment duration in our study population allows translation of the results in daily practice.

So far, studies have failed to consistently identify a single biomarker that can predict individual treatment response after starting or tapering TNFi with sufficient predictive value to be used in the individual patient with RA.\(^{29}\)\(^{30}\) We presume that prediction of clinical response to treatment is difficult due to the complex pathobiology of RA and the potential unknown effects of TNFi’s. As a consequence, measuring random laboratory markers is associated with a low priori chance of finding a biomarker. Furthermore, clinical response (based on changes in the DAS28 compared with baseline) is not solely the result of inflammatory RA activity but can be confounded by many other factors (eg, psychosocial) troubling prediction by a biomarker.

In conclusion, serum calprotectin has some predictive value for clinical response after starting anti-TNF treatment in RA, but it has no added value to other clinical factors such as DAS28-CRP. In patients with low disease activity, serum calprotectin was not predictive for clinical response after tapering anti-TNF treatment.

Author affiliations

1Department of Rheumatology, Sint Maartenskliniek, Nijmegen, The Netherlands
2Department of Internal Medicine, Radboud University Medical Center, Nijmegen, The Netherlands
3Department of Experimental Rheumatology, Radboud University Medical Center, Nijmegen, The Netherlands
4Department of Immunology, University of Münster, Münster, Germany
5Department of Rheumatology, Radboud University Medical Center, Nijmegen, The Netherlands

REFERENCES

8. Nordal HH, Brun JG, Hordvik M, et al. Calprotectin (S100A8/A9) and S100A12 are associated with measures of disease activity in a


