



Universiteit
Leiden
The Netherlands

The value of rheumatoid arthritis autoantibodies in disease pathogenesis and treatment prognosis

Moel, E.C. de

Citation

Moel, E. C. de. (2026, April 7). *The value of rheumatoid arthritis autoantibodies in disease pathogenesis and treatment prognosis*. Retrieved from <https://hdl.handle.net/1887/4302760>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4302760>

Note: To cite this publication please use the final published version (if applicable).



CHAPTER 6

Circulating calprotectin (S100A8/A9) is higher in rheumatoid arthritis patients that relapse within 12 months of tapering anti-rheumatic drugs

Emma C. de Moel¹, Jürgen Rech², Michael Mahler³, Johannes Roth⁴, Thomas Vogl⁴, Anne Schouffoer^{1,5}, Robbert J. Goekoop⁵, Tom W. J. Huizinga¹, Cornelia F. Allaart¹, René E. M. Toes¹, Georg Schett², Diane van der Woude¹

¹ Leiden University Medical Center, Leiden, the Netherlands

² Friedrich Alexander University of Erlangen-Nuremberg and Universitätsklinikum Erlangen, Germany

³ Inova Diagnostics, San Diego, California, United States

⁴ University of Muenster, Muenster, Germany

⁵ Haga Hospital, The Hague, the Netherlands

*Arthritis research & therapy. 2019;21(1):268.
doi: 10.1186/s13075-019-2064-y*

ABSTRACT

Objective

To investigate whether calprotectin (S100A8/A9 or MRP8/14), an inflammatory complex released by monocytes, could indicate residual subclinical inflammation in rheumatoid arthritis (RA) patients who are in stable remission on disease-modifying anti-rheumatic drugs (DMARDs) and serve as a marker for disease flare after DMARD tapering.

Methods

We used data from two trials. Patients from the IMPROVED study had early (< 2 years) RA, and when they achieved disease activity score remission (DAS44 < 1.6), they stopped methotrexate to attempt drug-free remission. Patients from the RETRO study had established RA in stable remission (DAS28 < 2.6) and either tapered by 50% or stopped (biological or conventional) DMARDs. Circulating calprotectin at the tapering time point was determined by ELISA, and its predictive value for flare (loss of remission) within 12 months of DMARD tapering/stopping was determined.

Results

In both IMPROVED (n = 104) and RETRO (n = 57), patients that flared within 12 months had higher calprotectin at the moment of DMARD tapering/stopping. Twofold higher calprotectin at the moment of DMARD tapering/stopping was associated with an increased risk (odds ratio) of flare of 1.07 (95% CI 0.98–1.18, p = 0.14) in the IMPROVED and 3.62 (95% CI 1.76–7.46, p < 0.001) in the RETRO. Correcting for clinical predictors of flare (DAS at study inclusion, anti-CCP2 positivity, gender) did not change these estimates. The area under the receiver operating curve of calprotectin levels for predicting flare within 12 months was 0.63 (95% CIs 0.51–0.76) in the IMPROVED study and 0.80 (95% CIs 0.69 to 0.92) in the RETRO study.

Conclusion

Circulating calprotectin levels in RA patients in remission on DMARDs are higher in patients that will flare upon DMARD tapering/stopping. Since the differences between the cohorts precluded definitive conclusions, more research is needed to determine whether calprotectin has prognostic value in predicting flare after attempting drug tapering in RA.

Trial registration

IMPROVED, ISRCTN11916566. RETRO, 2009-015740-42.

INTRODUCTION

In rheumatoid arthritis (RA), improvements in the control of inflammation by conventional synthetic and biologic disease-modifying antirheumatic drugs (DMARDs) are slowly shifting therapeutic goals from preventing damage by suppressing disease activity to tapering DMARDs in patients in stable remission to achieve drug-free remission (DFR). However, DFR will not be reached by all RA patients in remission, and the disease flare that can occur after unsuccessful DMARD tapering can increase disease burden that could have been avoided if the DMARDs had not been interrupted. Hence, identifying factors associated with increased risk of disease flare after tapering DMARDs would aid clinical decision-making.

Unfortunately, this endeavor has proven difficult. Most studies agree that seropositivity [both rheumatoid factor (RF) and anti-cyclic citrullinated protein antibodies (ACPA)], shared epitope presence, high disease activity at baseline, and long symptom duration are risk factors for flaring upon DMARD tapering (1, 2). Residual subclinical inflammation has also been proposed to be a major predictor of flare, prompting multiple studies to identify biomarkers sensitive in detecting residual inflammation at the moment of tapering, including multibiomarker disease activity scores (3) and inflammation on imaging (1). However, these markers are limited by variable associations and lack of replication.

More recently, circulating calprotectin has gained traction as a promising biomarker for RA disease activity. Calprotectin, also called S100A8/A9 or myeloid-related protein (MRP) 8/14, is a heterodimeric complex of the S100 family constitutively expressed in leukocytes. During infectious and inflammatory events, calprotectin acts as alarmin and amplifies inflammation by binding to receptors for advanced glycation end-products and Toll-like receptor 4. Traditionally measured in feces in inflammatory bowel disease diagnostics, calprotectin differs from hepatocyte-dependent acute-phase reactants like C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) because it is locally released by leukocytes at the sites of inflammation (4). In arthritis, activated granulocytes and tissue-resident macrophages in inflamed joints may contribute to calprotectin release. Various studies indicate that calprotectin actively promotes inflammation during arthritis(5, 6), is elevated in the synovial fluid and blood of RA patients, correlates consistently with markers of RA disease activity, is associated with ultrasound-detected inflammation, predicts future structural damage, and may have value in monitoring early response to biological DMARDs(7, 8). Furthermore, in juvenile idiopathic arthritis (JIA), a high level of calprotectin correlated with a high risk of relapse after etanercept(9) and methotrexate(10) discontinuation. The same was true in systemic-onset JIA after anakinra discontinuation (11).

All of these findings point to the intimate link between calprotectin, neutrophil activation, and inflammation in arthritis. If calprotectin indeed indicates residual subclinical

inflammation in RA patients who are in stable remission on DMARDs, it could serve as a practical biomarker for future disease flare after DMARD tapering and facilitate risk stratification to predict whether DMARD tapering will be successful. Such data could aid in clinical decision-making and replace the trial-and-error strategies by which DFR is currently—and infrequently—achieved. To that end, we investigated in two independent cohorts whether circulating calprotectin in patients who are in remission on DMARDs is associated with risk of disease flare after DMARD tapering and discontinuation.

METHODS

Study design, patient population, and outcomes

We used data from two multicenter randomized controlled trials: the IMPROVED study (an acronym for Induction therapy with MTX and Prednisone in Rheumatoid Or Very Early arthritic Disease) and the RETRO study (an acronym for Study of Reduction of Therapy in Patients with Rheumatoid Arthritis in Ongoing Remission). The IMPROVED study (12) enrolled 610 patients with early (< 2 years) untreated RA or undifferentiated arthritis and was steered at disease activity score remission ($\text{DAS44} < 1.6$). In patients achieving remission, DMARDs were stopped, aiming for drug-free remission (DFR), with treatment adjusted every 4 months according to whether remission was maintained. Initial treatment comprised MTX and high-dose prednisone. Subjects selected for this study were all 104 patients who (1) fulfilled the 2010 ACR/EULAR RA criteria, (2) achieved remission at 4 months and stopped MTX at 8 months, and (3) had serum available for calprotectin testing at 8 months.

The RETRO study(13) enrolled patients with established RA who had been in stable remission ($\text{DAS28-ESR} \leq 2.6$) for at least 6 months on conventional and/or biological DMARD treatment regimen. Patients were randomized to three treatment arms: (1) continue DMARDs at full dose for 12 months, (2) taper current dose by 50% for the next 12 months, or (3) reduce current dose by 50% for the first 6 months before entirely stopping all DMARDs. All patients in the taper arm (arm 2, $n = 33$) or stop arm (arm 3, $n = 24$) were included for this analysis.

Autoantibody status of the patients from both studies was determined by routine clinical testing for anti-CCP2 (IgG) and RF (IgM). The main outcome for both studies was flare within 12 months of tapering/stopping DMARDs, defined in the IMPROVED as a $\text{DAS44} \geq 1.6$ determined at 4-month intervals and in the RETRO as $\text{DAS28-ESR} > 2.6$ determined at 3-month intervals.

Calprotectin measurements

Patient serum was collected when patients were in remission and stored in -80°C until use. In the IMPROVED, circulating calprotectin was measured at baseline and 8 months (month of MTX stopping) using a fecal Calprotectin Extended Range kit adapted for

serum (QUANTA Lite®, Inova Diagnostics, research use only for serum). In the RETRO, calprotectin was measured in baseline samples, which was at study inclusion directly preceding DMARD tapering, before any tapering or stopping changes were made to existing DMARDs, using an in-house sandwich ELISA developed by Johannes Roth (Muenster) for predicting flares in juvenile idiopathic arthritis as previously described (11). Absorbance was converted to nanograms per milliliter using a standard curve of serial dilutions of a known concentration.

Statistical analysis

Calprotectin levels (in ng/mL) were compared between the groups using Mann-Whitney U tests, and correlations with inflammatory parameters were explored cross-sectionally using Spearman tests. Univariable and multivariable binary logistic regression was performed with disease flare within 12 months as the outcome. Calprotectin levels were log2-transformed and included as a continuous covariate. Area under the receiver operating characteristic curves (AUCs), test characteristics, and predictive values of calprotectin levels for flare were calculated, and the optimum cut-off was determined using the Liu index(14).

Additionally, logistic regression models were used to evaluate whether the addition of calprotectin to a clinical predictor model for flare significantly improved prediction of flare within 12 months. Clinical predictors of flare with a p value of ≤ 0.1 in univariable models were considered for the multivariable models to discover clinical predictors common to both cohorts, as well as specific to each cohort (as a sensitivity analysis, see Supplementary Table S1A-B). Areas under the ROC curve were calculated from the predicted values of the model with and without calprotectin, and compared using a Chi-square-based test for equality (15). All analyses were conducted separately for each cohort, using Stata SE 14.1.

RESULTS

In patients in remission, calprotectin is not associated with inflammatory parameters or autoantibody status

The characteristics of both cohorts are displayed in **TABLE 1**.

Circulating calprotectin levels, measured at the moment of DMARD tapering (for RETRO) or stop (for IMPROVED), were equal in the IMPROVED and the RETRO ($p = 0.51$), as well as between the two arms of the RETRO study ($p = 0.43$). Calprotectin levels were not correlated with acute-phase reactants (ESR, CRP), disease activity scores (DAS), or physical function as measured by the Health Assessment Questionnaire (HAQ) in either cohort, most likely due to the low variability of these parameters in these patients in remission (**SUPPLEMENTARY FIGURE S1**), and baseline demographics did not differ

between patients with high vs. low calprotectin levels (**SUPPLEMENTARY TABLE S2**). In baseline serum of the IMPROVED study, when patients had active disease activity ($\text{DAS44} \geq 2.6$) and had not yet received DMARD treatment, calprotectin did correlate with these inflammatory parameters (not shown), in a similar magnitude to what has been reported (7).

TABLE 1 Patient characteristics

	IMPROVED (N=104)	RETRO (N=57)
Age, mean years (SD)	49 (13)	55 (13)
Female, n (%)	67 (64%)	37 (65%)
BMI, mean (SD)	25 (4)	25 (4)
Ever smoker, n (%)	47 (46%)	18 (32%)
Disease duration, median (IQR)	19 (9-36) weeks [^]	5 (3-10) years
Anti-CCP2 IgG positive, n (%)	85 (83%)	34 (62%)
RF IgM positive, n (%)	79 (79%)	39 (68%)
Current DMARD use[†]		
Methotrexate	104 (100%)	47 (82%)
Glucocorticoids	0 (0%)	11 (19%)
Other csDMARDS	0 (0%)	6 [‡] (11%)
Biological DMARDs	0 (0%)	21 (37%)
Etanercept	-	5 (9%)
Adalimumab	-	6 (11%)
Tocilizumab	-	5 (9%)
Golimumab	-	2 (4%)
Certolizumab	-	3 (5%)
DAS44 at tapering/stop moment, mean (SD)	0.9 (0.4)	-
DAS28-ESR at tapering/stop moment, mean (SD)	-	1.8 (0.7)
Disease flare within 12 months	78 (75%)	26 (46%)
Calprotectin ng/mL, median (IQR)	1000 (230-2422)	1000 (650-2000)

Footnote: SD standard deviation, IQR interquartile range, Anti-CCP2 anti-citrullinated protein 2 antibody, RF rheumatoid factor, cs/bDMARD conventional synthetic/biological disease-modifying antirheumatic drugs, DAS disease activity score, ESR erythrocyte sedimentation rate. [^]Disease duration is based on the moment of study inclusion; in the IMPROVED study, this excludes the first 4 months of treatment [†]Categories are not mutually exclusive [‡]Patients received sulfasalazine (n = 2), hydroxychloroquine (n = 2), leflunomide (n = 1), or azathioprine (n = 1)

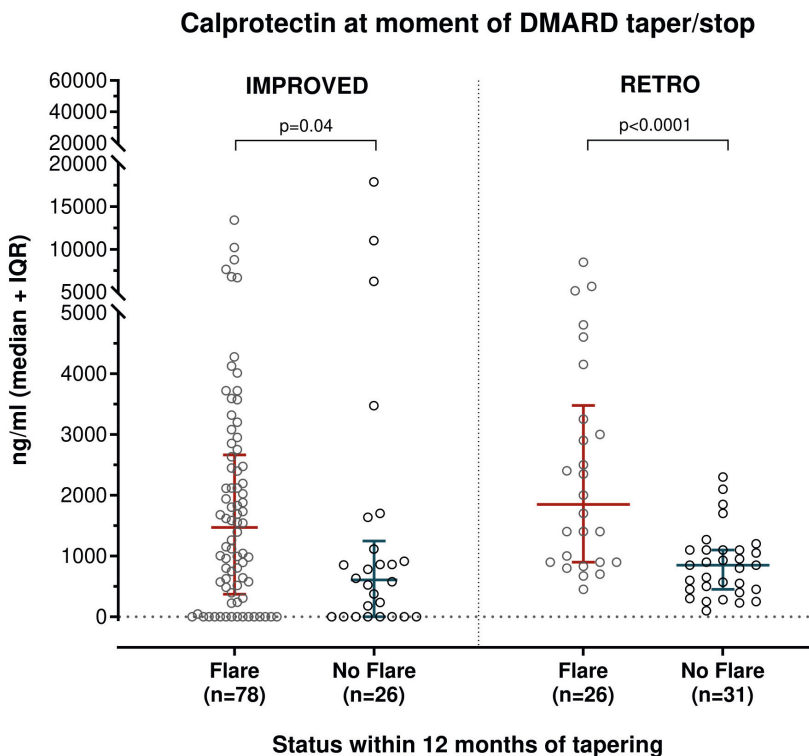
Calprotectin levels were higher in seropositive than seronegative patients, although not significantly: median (IQR) levels in anti-CCP2 IgG-positive vs. anti-CCP2 IgG-negative patients were 1115 (309–2749) ng/mL vs. 764 (0–1636) ng/mL in IMPROVED ($p = 0.21$)

and 1075 (800–2400) ng/mL vs. 930 (570–1850) ng/mL in RETRO ($p = 0.45$). In RF IgM-positive vs. RF IgM-negative patients, this was 1115 (237–2448) ng/mL vs. 982 (224–3318) ng/mL in IMPROVED ($p = 0.96$) and 1050 (550–1700) ng/mL vs. 965 (650–2300) ng/mL in RETRO ($p = 0.75$).

Calprotectin levels are higher preceding flare than no flare

In the IMPROVED study, disease flare was defined as DAS44 ≥ 1.6 within 12 months of stopping MTX. In the RETRO, flare was defined as DAS28-ESR > 2.6 within 12 months of tapering or stopping (biological or conventional) DMARDs. As indicated by **FIGURE 1**, calprotectin was significantly higher at the moment of DMARD tapering/stopping in patients that experienced a disease flare within 12 months.

FIGURE 1: Circulating calprotectin levels (ng/mL) at the moment of DMARD-tapering/stop, separated by whether patients experienced a disease flare within 12 months of tapering. p-values were calculated by Mann-Whitney-U tests

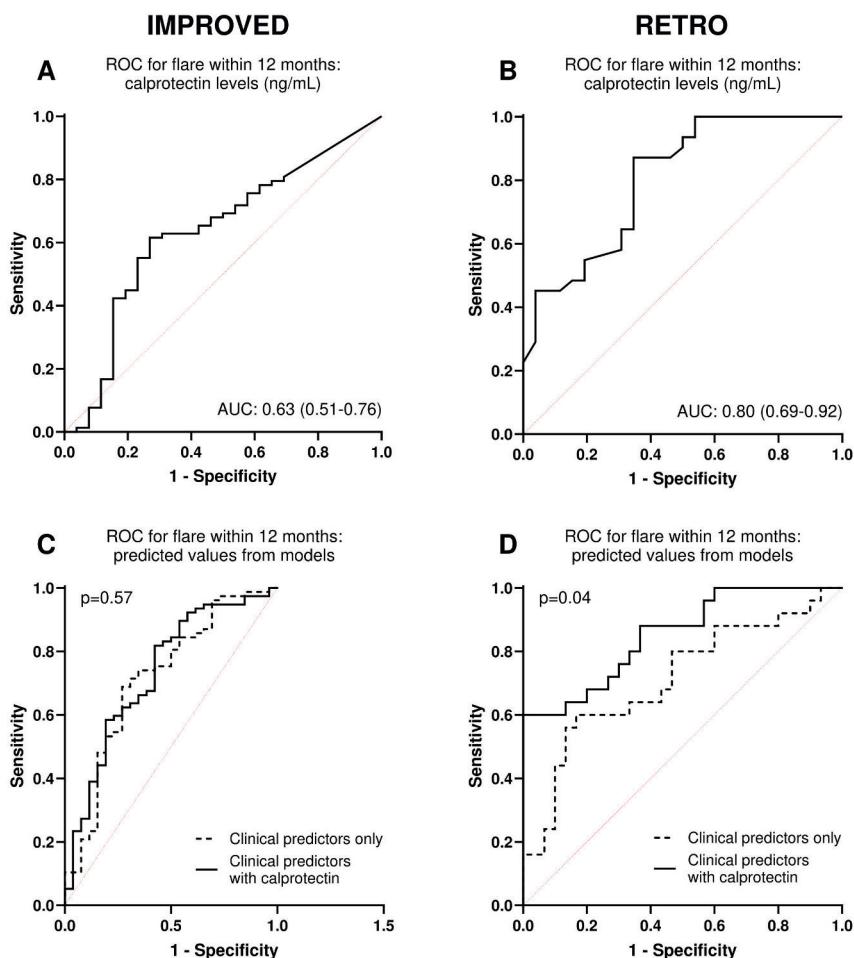


Calprotectin levels are associated with disease flare

In a logistic regression model with log₂-transformed calprotectin as a predictor, twofold higher calprotectin at the moment of DMARD tapering/stopping was associated with an increased risk (odds ratio) of flare of 1.07 (95% CI 0.98–1.18, $p = 0.14$) in the IMPROVED

study and 3.62 (95% CI 1.76–7.46, $p < 0.001$) in the RETRO study. Correcting for clinical predictors of flare (DAS at study inclusion, anti-CCP2 positivity, gender) resulted in similar estimates: 1.09 (95% CI 0.98–1.21, $p = 0.11$) in the IMPROVED study and 3.77 (95% CI 1.74–8.18, $p < 0.001$) in the RETRO study. Correcting for clinical predictors of flare selected for their predictive value specific for each cohort (ESR at tapering moment, gender, anti-CCP2 positivity, and DAS at study inclusion in IMPROVED; SJC and anti-CCP2 positivity in RETRO) resulted in similar estimates (not shown).

FIGURE 2: ROC-curves indicating the predictive value of calprotectin and clinical predictors for disease flare within 12 months of tapering in the IMPROVED (A,C) and RETRO (B,D) studies. A,B: circulating calprotectin levels (ng/mL) at the moment of DMARD-tapering/stop; C,D: comparison of models including clinical predictors only (dashed line) and clinical predictors in combination with circulating calprotectin levels (solid line). The p -value is based on the test for equality of AUCs.



The AUC of calprotectin levels predicting flare in the IMPROVED study was 0.63 (95% CI 0.51–0.76, $p = 0.002$), indicating modest to poor discriminatory capacity (**FIGURE 2A**). In the RETRO, calprotectin had a slightly better AUC of 0.80 (95% CI 0.69–0.92, $p = 0.0001$) for predicting flare within 12 months (**FIGURE 2B**).

In the IMPROVED, addition of calprotectin levels did not significantly improve a model including only the clinical predictors (DAS as baseline, anti-CCP2 positivity, gender): the clinical predictor model had an AUC of 0.72 (95% CI 0.60–0.84), compared to a model combining calprotectin and clinical predictors that had an AUC of 0.73 (95% CI 0.62–0.85) (p for test of equality of AUCs = 0.57; **FIGURE 2C**). However, in the RETRO, the addition of calprotectin to a clinical predictor model significantly improved the AUC from 0.71 (95% CI 0.57–0.85) to 0.85 (95% CI 0.75–0.95) (p for the test of equality of AUCs = 0.04; **FIGURE 2D**). These results were the same when cohort-specific clinical predictors mentioned above were used (not shown).

Test characteristics based on optimal and clinically relevant cutoffs

At an optimal cutoff determined using the Liu index, calprotectin levels above 936 ng/mL in the IMPROVED had the following test characteristics (95% CI) for disease flare within 12 months: 62% (50–72%) sensitivity, 73% (52–88%) specificity, 87% (76–95%) positive predictive value (PPV), and 39% (25–54%) negative predictive value (NPV). In the RETRO, calprotectin levels above 1335 ng/mL had slightly better test characteristics (95% CI): 65% (44–83%) sensitivity, 87% (70–96%) specificity, 81% (58–95%) PPV, and 75% (59–88%) NPV.

Since identifying patients with high flare risk that should not taper DMARDs is particularly relevant, we also examined a practical cutoff that maximized specificity. For the IMPROVED, calprotectin levels above 3571 ng/mL (corresponding to the 85th percentile) were 88% (70–98%) specific for flaring. In the RETRO, calprotectin levels above 2350 ng/mL (corresponding to the 79th percentile) were 100% (89–100%) specific for flaring. Patients classified as at risk for flare in this way (i.e., above this optimal cutoff) did not have higher inflammatory parameters (ESR, CRP, DAS, VAS, or HAQ) than those predicted not to flare (i.e., below the cutoff; not shown). In clinical practice, one could therefore infer from very high calprotectin levels that a patient is likely to flare regardless of other inflammatory parameters and should not taper DMARDs.

DISCUSSION

In this study, we evaluated the role of circulating calprotectin in RA as a predictor for disease flare upon DMARD tapering and discontinuation in two independent cohorts. Higher circulating calprotectin levels were associated with significantly increased risk of flare within 12 months after DMARD tapering in the RETRO study and a small, non-significant effect in the IMPROVED study. In the RETRO study, calprotectin significantly improved the identification of patients that would flare within 12 months over clinical predictors alone.

Calprotectin has been reported to correlate better with clinical disease activity of RA than acute-phase reactants do (16). These findings suggest that calprotectin levels may reflect local inflammatory processes in the joints and therefore make it of particular value in predicting whether inflammation has been abrogated and patients can achieve drug-free remission. Possessing a marker that can accurately indicate residual inflammation in patients in remission would help to optimize therapeutic decision-making by accurately stratifying which patients should and should not attempt to taper DMARDs.

In line with the literature (16), circulating calprotectin levels in our study were variable and sometimes quite high, even in patients in remission, a disease state in which systemic inflammation is, in principle, low. Although we were not able to replicate previous findings regarding the correlation between calprotectin and joint indices due to the lack of variability in joint indices, the fact that calprotectin did not correlate with other inflammatory parameters in these patients in remission suggests that calprotectin may indeed indicate residual inflammation not measured by these parameters, probably because it is highly expressed and released specifically at local sites of inflammation (4).

However, calprotectin's value for predicting disease flare was not consistent in both cohorts. Although the direction of the found effect was the same in both cohorts, the effect sizes we found are quite divergent and only the RETRO study reached consistent statistical significance, possibly due to the differences between the cohorts. One major difference is the use of DAS44 in the IMPROVED and DAS28-ESR in the RETRO, which does not detect affected joints in the feet and thus could miss peripheral disease activity in the feet. It also is conceivable that calprotectin is not sensitive enough to detect residual inflammation when DAS is extremely low, as was the case in the IMPROVED study ($\text{DAS44} < 1.6$ instead of $\text{DAS28-ESR} < 2.6$ in RETRO). However, a sensitivity analysis restricting the analyses in the IMPROVED cohort only to patients on the higher side of the spectrum ($\text{DAS44} \geq 0.9$ at 8 months) did not indicate that calprotectin had more value in patients with slightly higher disease activity (not shown). It is also possible that calprotectin cannot predict minute changes in DAS (e.g., small flares), but a sensitivity analysis in the IMPROVED study in which $\text{DAS44} \geq 2.6$ within 12 months was regarded as flare (more comparable to the RETRO definition) did not indicate that calprotectin had greater value in predicting larger DAS44 change-based flares than smaller flares (not shown). Differences in the biological use, disease durations, and remission duration could explain the discrepant results as well. However, stratification on biological use in the RETRO resulted in very small numbers and inconclusive results, and the disease and remission durations between the cohorts were so divergent that stratification on this parameter would not yield fair comparisons.

Some previous studies investigating calprotectin and flares in rheumatic diseases other than RA support calprotectin as a marker for flare after tapering DMARDs (10, 11), but Tweehuysen et al. found no value of calprotectin for predicting successful DMARD

tapering in RA (17). The main difference between our cohorts and the DRESS study (an acronym for Dose REDuction Strategy of Subcutaneous TNF inhibitors) they investigated was that all patients in their study tapered biological DMARDs, specifically TNF α inhibitors. It is possible that TNF α inhibitors may directly inhibit the release of calprotectin in vivo and dampen calprotectin's value as a marker of local inflammation, thereby also limiting its ability to predict future disease flare. However, beyond the upregulation of calprotectin by TNF α in murine models (18, 19), an inverse correlation between trough serum TNF α -inhibitor levels and circulating calprotectin (20), a consistent decrease of calprotectin upon TNF α -inhibitor initiation(8), no studies to date have explored this possibility.

Surprisingly, the population of the RETRO study, in which the association of calprotectin with flare was most convincing, is more similar to the DRESS study (long-standing disease, biological use including TNF α inhibitors, higher DAS at tapering) than the IMPROVED population is, even though the latter two cohorts showed little to no association of calprotectin with flare. This makes it very difficult to speculate on the reasons why these three cohorts show such divergent results, as there is clearly no pattern regarding patient characteristics and strength of associations found. Because of this, our results cannot definitively refute or commend calprotectin as a marker for successful DMARD tapering.

A further limitation to consider is that calprotectin was determined using different assays, and although circulating concentrations were converted to nanograms per milliliter using a known concentration curve in both cohorts, small differences in the sensitivity of the assays could remain. Sample handling may also have differed between various centers participating in the studies. However, there was no difference in calprotectin levels between the cohorts on the group level, so the effect of these limitations is most likely minor.

Although our results show modest effects, they are important because they situate calprotectin in a new field of prognostication in RA—that of drug-free remission. Currently, there is no consensus regarding calprotectin's use as a measure of disease activity or a biomarker for treatment response, let alone the possibility of using it as a marker of future disease flare, and more studies are needed to evaluate calprotectin in this role.

CONCLUSION

In conclusion, we found that calprotectin levels are higher in RA patients that will flare upon DMARD tapering. More research is needed to validate calprotectin as a biomarker for flare in RA patients tapering or discontinuing DMARDs.

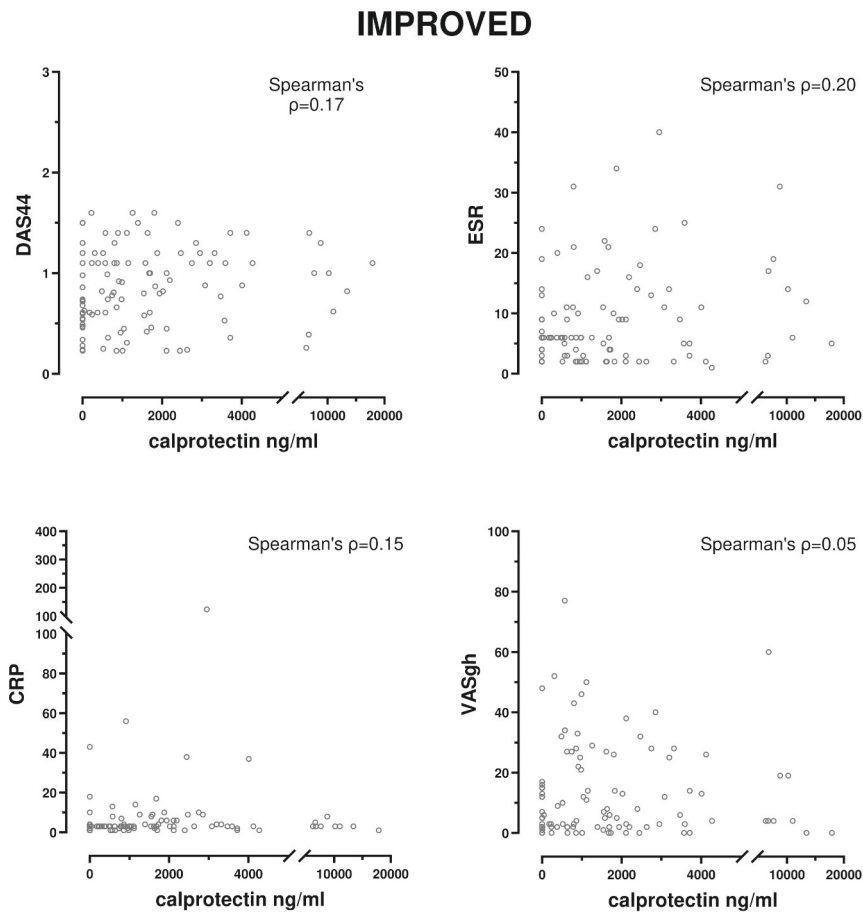
REFERENCES

- Schett G, Emery P, Tanaka Y, Burmester G, Pisetsky DS, Naredo E, et al. Tapering biologic and conventional DMARD therapy in rheumatoid arthritis: current evidence and future directions. *Annals of the rheumatic diseases*. 2016;75(8):1428-37.
- van den Broek M, Huizinga TW, Dijkmans BA, Allaart CF. Drug-free remission: is it already possible? Current opinion in rheumatology. 2011;23(3):266-72.
- Rech J, Hueber AJ, Finzel S, Englbrecht M, Haschka J, Manger B, et al. Prediction of disease relapses by multibiomarker disease activity and autoantibody status in patients with rheumatoid arthritis on tapering DMARD treatment. *Annals of the rheumatic diseases*. 2016;75(9):1637-44.
- Vogl T, Eisenblätter M, Voller T, Zenker S, Hermann S, van Lent P, et al. Alarmin S100A8/S100A9 as a biomarker for molecular imaging of local inflammatory activity. *Nat Commun*. 2014;5:4593.
- van Lent PL, Grevers L, Blom AB, Sloetjes A, Mort JS, Vogl T, et al. Myeloid-related proteins S100A8/S100A9 regulate joint inflammation and cartilage destruction during antigen-induced arthritis. *Annals of the rheumatic diseases*. 2008;67(12):1750-8.
- Vogl T, Stratis A, Wixler V, Voller T, Thurainayagam S, Jorch SK, et al. Autoinhibitory regulation of S100A8/S100A9 alarmin activity locally restricts sterile inflammation. *The Journal of clinical investigation*. 2018;128(5):1852-66.
- Bae SC, Lee YH. Calprotectin levels in rheumatoid arthritis and their correlation with disease activity: a meta-analysis. *Postgrad Med*. 2017;129(5):531-7.
- Abildtrup M, Kingsley GH, Scott DL. Calprotectin as a biomarker for rheumatoid arthritis: a systematic review. *The Journal of rheumatology*. 2015;42(5):760-70.
- Anink J, Van Suijlekom-Smit LW, Otten MH, Prince FH, van Rossum MA, Dolman KM, et al. MRP8/14 serum levels as a predictor of response to starting and stopping anti-TNF treatment in juvenile idiopathic arthritis. *Arthritis research & therapy*. 2015;17:200.
- Foell D, Wulffraat N, Wedderburn LR, Wittkowski H, Frosch M, Gerss J, et al. Methotrexate withdrawal at 6 vs 12 months in juvenile idiopathic arthritis in remission: a randomized clinical trial. *Jama*. 2010;303(13):1266-73.
- Holzinger D, Frosch M, Kastrop A, Prince FH, Otten MH, Van Suijlekom-Smit LW, et al. The Toll-like receptor 4 agonist MRP8/14 protein complex is a sensitive indicator for disease activity and predicts relapses in systemic-onset juvenile idiopathic arthritis. *Annals of the rheumatic diseases*. 2012;71(6):974-80.
- Wevers-de Boer K, Visser K, Heimans L, Roday HK, Molenaar E, Groenendaal JH, et al. Remission induction therapy with methotrexate and prednisone in patients with early rheumatoid and undifferentiated arthritis (the IMPROVED study). *Annals of the rheumatic diseases*. 2012;71(9):1472-7.
- Haschka J, Englbrecht M, Hueber AJ, Manger B, Kleyer A, Reiser M, et al. Relapse rates in patients with rheumatoid arthritis in stable remission tapering or stopping antirheumatic therapy: interim results from the prospective randomised controlled RETRO study. *Annals of the rheumatic diseases*. 2016;75(1):45-51.
- Liu X. Classification accuracy and cut point selection. *Stat Med*. 2012;31(23):2676-86.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-45.

16. Inciarte-Mundo J, Ruiz-Esquide V, Hernandez MV, Canete JD, Cabrera-Villalba SR, Ramirez J, et al. Calprotectin more accurately discriminates the disease status of rheumatoid arthritis patients receiving tocilizumab than acute phase reactants. *Rheumatology*. 2015;54(12):2239-43.
17. Tweehuysen L, den Broeder N, van Herwaarden N, Joosten LAB, van Lent PL, Vogl T, et al. Predictive value of serum calprotectin (S100A8/A9) for clinical response after starting or tapering anti-TNF treatment in patients with rheumatoid arthritis. *RMD Open*. 2018;4(1):e000654.
18. Xu K, Geczy CL. IFN-gamma and TNF regulate macrophage expression of the chemotactic S100 protein S100A8. *J Immunol*. 2000;164(9):4916-23.
19. Koenders MI, Marijnissen RJ, Devesa I, Lubberts E, Joosten LA, Roth J, et al. Tumor necrosis factor-interleukin-17 interplay induces S100A8, interleukin-1beta, and matrix metalloproteinases, and drives irreversible cartilage destruction in murine arthritis: rationale for combination treatment during arthritis. *Arthritis and rheumatism*. 2011;63(8):2329-39.
20. Inciarte-Mundo J, Ramirez J, Hernandez MV, Ruiz-Esquide V, Cuervo A, Cabrera-Villalba SR, et al. Calprotectin and TNF trough serum levels identify power Doppler ultrasound synovitis in rheumatoid arthritis and psoriatic arthritis patients in remission or with low disease activity. *Arthritis research & therapy*. 2016;18(1):160.

SUPPLEMENTARY FIGURES

FIGURE S1: Circulating calprotectin levels (ng/mL) at the moment of DMARD tapering/stop does not correlate with inflammatory parameters in the IMPROVED (left) or the RETRO (right) study. VASgh = VAS global health.



RETRO

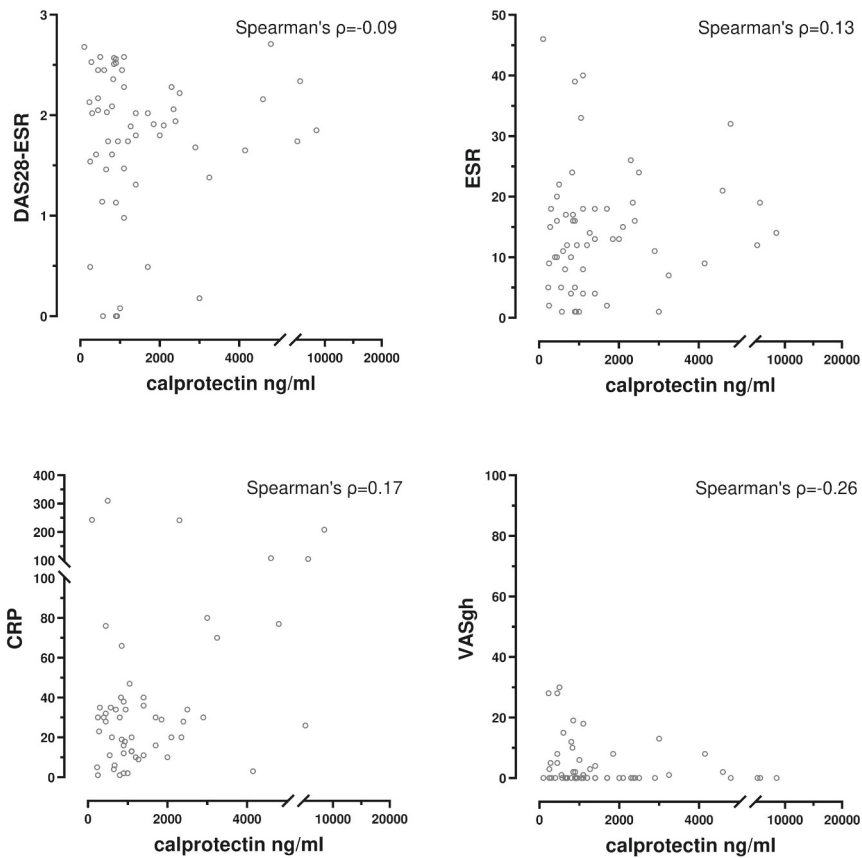


TABLE S1A: Model building for outcome flare in the IMPROVED study. All variables were tested univariably for association with the outcome; only those variables with $p \leq 0.1$ were included in multivariable models. Multivariable models were tested separately for activity parameters (top) and for demographic/baseline variables (bottom).

IMPROVED study (n=104)							
Activity parameters at tapering/stop moment	Flare (n=78)	No Flare (n=26)	Univariable OR (95% CI)	p-value	Multivariable OR (95% CI)	p-value	
DAS44, mean (SD)	0.9 (0.4)	0.7 (0.4)	3.67 (1.07-12.6)	0.04	1.16 (0.21-6.44)	0.87	
ESR, median (IQR)	6 (3-14)	6 (2-9)	1.10 (1.01-1.19)	0.04	1.10 (1.01-1.20)	0.04	
CRP, median (IQR)	3 (3-6)	3 (1-3)	1.01 (0.97-1.06)	0.56			
VASgh, median (IQR)	10 (2-26)	4 (2-8)	1.05 (1.00-1.09)	0.03	1.03 (0.98-1.08)	0.21	
Tender joint count-28, median (IQR), (minimum-maximum)	0 (0-0), (0-4)	0 (0-0), (0-4)	1.02 (0.62-1.70)	0.93			
Swollen joint count-28, median (IQR), (minimum-maximum)	0 (0-0), (0-3)	0 (0-0), (0-1)	1.84 (0.47-7.28)	0.87			
HAQ, median (IQR)	0 (0-0.38)	0 (0-0.13)	33.0 (1.53-711)	0.03	18.8 (0.71-498)	0.08	

TABLE S1A: Model building for outcome flare in the IMPROVED study. All variables were tested univariably for association with the outcome; only those variables with $p \leq 0.1$ were included in multivariable models. Multivariable models were tested separately for activity parameters (top) and for demographic/baseline variables (bottom). (continued)

Demographic/baseline variables	Flare (n=78)	No Flare (n=26)	Univariable OR (95% CI)	p-value	Multivariable OR (95% CI)	p-value
Age, mean years (SD)	49 (14)	51 (10)	0.99 (0.95-1.02)	0.43		
Female, %	69%	50%	2.25 (0.91-5.57)	0.08	3.10 (1.11-8.66)	0.03
BMI, mean (SD)	26 (4)	25 (3)	1.06 (0.92-1.22)	0.41*	1.15 (0.97-1.36)	0.11
Ever smoker, %	44%	52%	0.71 (0.29-1.76)	0.46		
Disease duration, weeks median (IQR)	19 (3-39)	18 (11-26)	1.02 (0.99-1.04)	0.14		
Anti-CCP2 IgG positive, %	84%	77%	1.63 (0.54-4.89)	0.39*	3.47 (0.95-12.6)	0.05
RF IgM positive, %	80%	75%	1.36 (0.46-4.00)	0.58		
DAS44, mean (SD)	3.1 (0.8)	2.7 (0.9)	2.11 (1.12-3.97)	0.02	2.45 (1.24-4.85)	0.01
HAQ, median (IQR)	1 (0.63-1.5)	0.75 (0.38-1.13)	1.95 (0.89-4.26)	0.10		

* Anti-CCP2 IgG positivity and BMI were noteworthy at $p=0.07$ and $p=0.03$ respectively as a likely predictors for flare in a screening model including all demographic/baseline variables together (not shown), and were therefore included in the subsequent multivariable model. Final IMPROVED-specific multivariable model including ESR at tapering moment, gender, anti-CCP2 positivity, and DAS44 at baseline had a pseudo-R² of 14%. Final model including predictors common to both cohorts (DAS at baseline, anti-CCP2 positivity, gender) had a pseudo R² of 11% in the IMPROVED study. Choice of these common predictors was based on known predictors of flare reported in literature as well as commonalities in predictors for both cohorts, reported above.

TABLE S1B: Model building for outcome flare in the RETRO study. All variables were tested univariably for association with the outcome; only those variables with $p \leq 0.1$ were included in multivariable models. Multivariable models were tested separately for activity parameters (top) and for demographic/baseline variables (bottom).

RETRO study (n=57)						
Activity parameters at tapering/stop moment	Flare (n=26)	No Flare (n=31)	Univariable OR (95% CI)	p-value	Multivariable OR (95% CI)	p-value
DAS28-ESR at tapering/stop moment, mean (SD)	1.8 (0.7)	1.8 (0.8)	1.02 (0.50-2.11)	0.95		
ESR, median (IQR)	13 (7-19)	13 (8-18)	1.00 (0.94-1.05)	0.87		
CRP, median (IQR)	32 (11-70)	20 (13-34)	0.94 (0.41-2.20)	0.89		
VASgh, median (IQR)	0 (0-5)	0 (0-4)	0.97 (0.90-10.5)	0.47		
Tender joint count-28, median (IQR), (minimum-maximum)	0 (0-0), (0-1)	0 (0-0), (0-1)	0.27 (0.03-2.58)	0.26		
Swollen joint count-28, median (IQR), (minimum-maximum)	0 (0-0), (0-2)	0 (0-0), (0-1)	6.27 (0.74-53.18)	0.09	n.d.*	

TABLE S1B: Model building for outcome flare in the RETRO study. All variables were tested univariably for association with the outcome; only those variables with $p \leq 0.1$ were included in multivariable models. Multivariable models were tested separately for activity parameters (top) and for demographic/baseline variables (bottom). (continued)

Demographic/baseline variables	Flare (n=26)	No Flare (n=31)	Univariable OR (95% CI)	p-value	Multivariable OR (95% CI)	p-value
Age, mean years (SD)	56 (13)	55 (13)	1.00 (0.96-1.04)	0.98		
Female, %	69%	61%	1.42 (0.47-4.28)	0.62		
BMI, mean (SD)	25 (3)	26 (4)	0.91 (0.79-1.05)	0.22		
Ever smoker, %	31%	32%	0.93 (0.30-2.87)	0.90		
Disease duration, years median (IQR)	8 (3-10)	4 (2-8)	1.04 (0.98-1.12)	0.20		
Anti-CCP2 IgG positive, %	80%	47%	4.57 (1.36-15.40)	0.01	n.d.*	
RF IgM positive, %	73%	65%	1.49 (0.48-4.65)	0.49		
Biological DMARDs (%)	38%	35%	1.14 (0.39-3.34)	0.82		
Randomization arm 3 (Taper) (%)	50%	35%	1.82 (0.63-5.27)	0.27		

* Multivariable model not done (n.d.) as only one variable within each model subsection was significantly associated with the outcome. Final RETRO-specific multivariable model including swollen joint count-28 and anti-CCP2 positivity had a pseudo-R² of 15%. Final model including predictors common to both cohorts (DAS at baseline, anti-CCP2 positivity, gender) had a pseudo R² of 9% in the RETRO study. Choice of these common predictors was based on known predictors of flare reported in literature as well as commonalities in predictors for both cohorts, reported above.

TABLE S2: Inflammatory parameters and baseline demographics separated by low/high circulating calprotectin levels (ng/mL) split on its median at the moment of DMARD tapering/stop. P-values are based on t-tests, Mann-Whitney tests, or Chi-square tests for the differences in mean, median, and frequency of reported variables, respectively.

IMPROVED study (n=104)	Low calprotectin (n=52)	High calprotectin (n=52)	p-value
DAS44 at tapering moment, mean (SD)	0.8 (0.4)	0.9 (0.4)	0.22
ESR, median (IQR)	6 (3-9)	8 (3-15)	0.38
CRP, median (IQR)	3 (2-3)	6 (3-8)	0.03
VASgh, median (IQR)	6 (2-25)	6 (2-19)	0.73
Tender joint count-28, median (IQR), (minimum-maximum)	0 (0-0), (0-3)	0 (0-0), (0-4)	0.49
Swollen joint count-28, median (IQR), (minimum-maximum)	0 (0-0), (0-1)	0 (0-0), (0-3)	0.53
Age, mean years (SD)	49 (14)	49 (12)	0.95
Female, %	62%	67%	0.54
BMI, mean (SD)	25 (3)	26 (4)	0.09
Ever smoker, %	42%	49%	0.49
Disease duration, weeks median (IQR)	19 (10-30)	19 (9-50)	0.63
Anti-CCP2 IgG positive, %	81%	84%	0.64
RF IgM positive, %	78%	80%	0.73

TABLE S2: Inflammatory parameters and baseline demographics separated by low/high circulating calprotectin levels (ng/mL) split on its median at the moment of DMARD tapering/stop. P-values are based on t-tests, Mann-Whitney tests, or Chi-square tests for the differences in mean, median, and frequency of reported variables, respectively. (continued)

RETRO study (n=57)	Low calprotectin (n=29)	High calprotectin (n=28)	p-value
DAS28-ESR at tapering/stop moment, mean (SD)	1.7 (0.9)	1.8 (0.6)	0.67
ESR, median (IQR)	11 (5-17)	14 (10-19)	0.16
CRP, median (IQR)	28 (11-35)	29 (13-59)	0.36
VASgh, median (IQR)	1 (0-8)	0 (0-1.5)	0.10
Tender joint count-28, median (IQR), (minimum-maximum)	0 (0-0), (0-1)	0 (0-0), (0-0)	0.02
Swollen joint count-28, median (IQR), (minimum-maximum)	0 (0-0), (0-2)	0 (0-0), (0-1)	1.00
Age, mean years (SD)	53 (12)	59 (5)	0.08
Female, %	66%	64%	0.92
BMI, mean (SD)	26 (5)	25 (3)	0.39
Ever smoker, %	38%	25%	0.29
Disease duration, years median (IQR)	5 (3-10)	6 (2-10)	0.72
Anti-CCP2 IgG positive, %	57%	66%	0.47
RF IgM positive, %	66%	71%	0.63
Biological DMARDs (%)	34%	43%	0.52
Randomization arm 3 (Stop) (%)	38%	46%	0.52