Study Title: Treatments Against RA and Effect on FDG-PET/CT (TARGET) Trial
Document Name: Statistical Analysis Plan, v14
Date: May 8, 2021
ClinicalTrials.gov Identifier: NCT02374021

Glossary of Abbreviations

MTX	Methotrexate
SSZ	Sulfasalazine
HCQ	Hydroxychloroquine
TNFi	TNF inhibitor
RA	Rheumatoid arthritis
MHDA	Moderate or high disease activity
LDAR	Low disease activity or remission
FDG PET/CT	FDG positron emission tomography/computed tomography
DMARD	Disease modifying anti-rheumatic drug
MTX-IR	Methotrexate inadequate responder
ROI	Region of interest
SUV	Standardized uptake value
TBR	Total to background ratio
MDS	Most diseased segment
ANCOVA	Analysis of covariance
SQ	Subcutaneous
TJC	Tender joint count
SJC	Swollen joint count
PtG	Patient global arthritis activity
CRP	C-reactive protein
BMI	Body mass index

I. AIMS

Aim 1. To compare the effects on vascular inflammation of TNFi + MTX versus triple therapy [MTX+SSZ+HCQ] in patients with RA who are inadequate responders to MTX monotherapy. *We hypothesize that TNFi +MTX will reduce vascular inflammation to a greater extent than triple therapy when measured at 6 months.*

Aim 2. To compare the effects on vascular inflammation of achieving low disease activity or remission (LDAR) vs remaining in moderate-high disease activity (MHDA). These pre-specified secondary analyses will pool the treatment arms to examine whether achievement of a disease activity target is associated with greater reduction in vascular inflammation.

2a. To use the Disease Activity Score-CRP-28 (DAS-28-CRP[1]) to categorize treatment response and correlate it with vascular inflammation. *We hypothesize that subjects in DAS-28-CRP defined LDAR at 6 months will have less vascular inflammation than those with persistent MHDA.*

2b. To use a multi-biomarker of RA disease activity (Vectra-DA) to categorize treatment response and correlate it with vascular inflammation. *We hypothesize that subjects in Vectra DA-defined LDAR at 6 months will have less vascular inflammation than those with persistent MHDA*. All 12 proteins in Vectra-DA are also recognized CVD biomarkers[2], thus, each one will also be analyzed alone and in combinations to determine whether a different algorithm of the Vectra-DA components may better correlate with vascular inflammation. We will also compare the correlations of the Vectra-DA and the DAS-28-CRP with vascular inflammation to determine if they are independent.

2c. In this exploratory sub-aim, we will use joint inflammation as measured by articular FDG PET/CT to categorize treatment response and correlate it with vascular inflammation. *We hypothesize that subjects in articular FDG defined LDAR at 6 months will have less vascular inflammation than those with persistent MHDA.* We will also compare the correlations of the articular FDG uptake and DAS-28-CRP with vascular inflammation to determine if they are independent.

II. STUDY POPULATION

Subjects for this trial all fulfilled ACR/EULAR criteria for RA and were deemed methotrexate-inadequate responders (MTX-IRs, DAS28>3.2) by their treating rheumatologist. They could not be current or recent (in the past 6 months) users of a biologic DMARD and could not have previously been deemed inadequate responders to a TNF antagonist. To enrich our trial subjects for vascular disease, we required patients to be at least 45 years old for men, and 50 years old for women. Detailed inclusion and exclusion criteria are listed in **Table 1.**

Table 1: Detailed Eligibility Criteria
(Subjects who meet all of the following criteria at Screening are eligible for enrollment into the study)

Inclusion

- Written informed consent signed by the subject;
- Fulfill ACR/EULAR 2010 criteria for RA;
- Men ≥ 45 years and women ≥ 50 years;
- MTX for ≥ 8 weeks at ≥ 15mg weekly or on at least 7.5mg of methotrexate weekly for ≥8 weeks with a documented intolerance of higher MTX doses, and stable dose for the previous 4 weeks;
- DAS28 score > 3.2;
- Able to swallow pills;
- Males and females with reproductive potential must agree to practice effective measures of birth control;
- If taking prednisone (or equivalent corticosteroid), the dose must be ≤ 10 mg/day at the time of the baseline FDG PET/CT scan and must NOT change by more than ±3.0 mg for the four weeks prior to the baseline FDG PET/CT (If subjects are taking steroids every other day, divide dose by 2 to evaluate eligibility.);
- If taking a low- or moderate-intensity statin, the dose must be stable for six weeks prior to screening and must not change during the six months of the trial; and
- Willing to comply with all study procedures and be available for the duration of the study
- Rheumatoid arthritis without psoriasis or with psoriasis if rheumatoid factor ≥2x ULN or anti-CCP ≥2x ULN.

Exclusion

- Use of biologic DMARD or small molecule DMARD (i.e. tofacitinib) in the past 6 months, use of Rituximab ever;
- Non-biologic DMARDs other than MTX or HCQ for two months prior to Screening
- Considered to be an etanercept (Enbrel) or adalimumab (Humira) failure by their primary rheumatologist;
- Current use of > 10mg per day of prednisone;
- Current use or use within the previous 12 months of a high-intensity statin (atorvastatin/Lipitor 40mg or higher, rosuvastatin/Crestor 10mg or higher) or PCSK9 inhibitor (Alirocumab/ Praluent, Evolocumab/Repatha, or Bococizumab);
- Prior patient reported, physician diagnosed clinical CV event: myocardial infarction, angina, stroke, uncompensated or severe heart failure (NYHA class III or IV), prior vascular procedure (coronary artery angioplasty or stenting, carotid endarterectomy, coronary artery bypass surgery);
- Demyelinating disease;
- Any of the following forms of arthritis that may otherwise explain the subject's RA symptoms: Psoriatic Arthritis, Reactive Arthritis, Juvenile Idiopathic Arthritis, Ankylosing Spondylitis, Polymyalgia Rheumatica;
- Any of the following other autoimmune and/or chronic inflammatory diseases: Inflammatory Bowel Disease, Crohn's disease, Cutaneous or Systemic Lupus, Systemic Vasculitis, Giant Cell Arteritis, Polymyositis, Dermatomyositis, Sarcoidosis, or Scleroderma ;
- Cancer treated in last five years (except basal and squamous cell) or any lymphoma or melanoma;
- Type I diabetes mellitus or Type II diabetes treated with insulin or uncontrolled with HbA1c ≥ 7% from the past 6 months;
- Known history of transient ischemic attack (TIA), stroke, myocardial infarction, or revascularization for coronary or peripheral artery disease;
- Known pregnancy, HIV, hepatitis B, hepatitis C, active (or untreated latent) TB;
- Known sulfa allergy or other known hypersensitivity to any of the trial agents or G6PD deficiency;
- Known macular disease or known retinal disease;
- Baseline blood count, renal or liver abnormalities as follows: WBC < 3.5 X1000 n/ul, Hematocrit < 30%, Platelet count < 90 x1000 n/ul, estimated glomerular filtration rate < 50 ml/min, AST (liver function test) > 60 U/L, ALT (liver function test) >84 U/L;
- Intra-articular injection of corticosteroids within the 4 weeks prior to the potential baseline FDG PET/CT; or
- Two or more of the following high dose radiation scans: CT scan with contrast, angiogram, SPECT nuclear medicine scan, myocardial (cardiac) perfusion scan in the past year.

III. INTERVENTIONS

Treatments. Patients assigned to the TNFi arm received either etanercept 50 mg SQ qwk or adalimumab 40 mg SQ every other week. At 12 weeks, if the patient had not achieved a EULAR good response[3] (DAS28 ≤ 3.2 AND improvement in DAS28 from baseline ≥ 1.2) the dose was increased (adalimumab 40 mg every other to every week) or the medication will be switched (etanercept to adalimumab 40 mg every other week). The patient remained on that medication/dose to the end of the study. Patients assigned to triple therapy began SSZ 1 gm bid and HCQ 200 mg bid (not to exceed 6.5mg/kg). At 6 weeks, if they have not achieved a good EULAR response, SSZ was increased to 1.5 gm bid. If a good EULAR response has not been achieved by 12 weeks, MTX was switched to a SQ route and/or the dose increased (if dose at enrollment is ≤ 25 mg/wk). At 18 weeks, if the patient was not in low disease activity or remission, the investigator could offer several possible modifications in treatment: for patients taking triple therapy, MTX would be substituted with leflunomide; for patients taking adalimumab every other week, the frequency could be increased to weekly; and for patients taking etanercept, they could switch to adalimumab.

Withdrawal Criteria. We conservatively assumed 15% withdrawal or cross-over and observed approximately 15% withdrawal and no cross-over. Withdrawal from the study was indicated for a cautionary laboratory value, a new malignancy other than basal or squamous cell, repeated subject non-compliance, or loss to follow-up. If study treatment was stopped (e.g., for an adverse event) and there is no safety issue precluding it, the patient underwent FDG PET/CT scan within 8 weeks of the last dose of randomized treatment.

IV. OUTCOMES

Primary outcome definition – change in vascular FDG uptake: Imaging was performed prior to randomization and at week 24 while on study drug using methods reported previously[4-8], and following detailed image acquisition guidelines.. In brief, FDG was administered (10 mCi) after an over-night fast. Imaging was performed 90 minutes later using a PET/CT scanner (with at least 8 CT detectors; most have 16) with the patient supine and head and chest adjusted to a standardized position by use of laser as outlined in the Imaging Protocol. Imaging was performed across 3-4 bed positions, covering the neck and chest. CT attenuation correction was performed using tube voltage of 120 kVp, and current of ~ 40mAs. PET data were acquired in 3-D mode over 10 minutes per bed position and stored on 256x256 matrix. Attenuation-corrected images were reconstructed yielding ~4 mm effective resolution. Following neck and chest scans, a whole-body PET/CT scan was performed for articular FDG uptake using the same voltage parameters. The arms were reoriented in an extended position with palms on anterior thighs.

Patient identifiers were removed by the imaging site investigators and transferred to the Imaging Core, using secure FTP. The system for image transfer has been validated through several prior multi-center trials. Each of the baseline and week 24 images was sent to the Imaging Core for quality control as soon as it is obtained. Upon receipt by the Imaging Core, images are assessed for protocol compliance and image quality, then transferred to a dedicated multi-modality workstation. Image analysis was performed by an experienced reader with paired attenuation image sets of the vessels of interest (right and left carotid and aorta), evaluated side-by-side with scrambled time-points and blinding to treatment as previously reported[9]. The target tissues were matched so that the same locations are measured for both time-points. Thereafter, regions of interest (ROIs) were drawn around the target vessel (in axial orientation) to provide maximum standardized uptake values (SUV) for each ROI. The SUV is the decay-corrected tissue concentration of FDG (in kBq/ml) divided by the injected dose per body weight (kBq/g). Drawing of ROIs was repeated along the length of the vessel (every ~3.5 mm along the long axis of the vessel) to provide a stack that compose the whole vessel[9]. Background (using the SVC blood pool) corrected maximum SUVs are averaged to provide a whole vessel target-to-background ratio (TBR). Absence of correction for partial volume effect has not been a limitation in assessment of larger vessels, such as the carotids and the aorta, as shown by published validation studies[10].

The primary PET/CT parameter used to evaluate vascular inflammation will be the <u>mean of the maximum TBR of the MDS of the index vessel (index vessel MDS TBRmeanmax),</u> based on work showing that this endpoint provides the most sensitive measure of treatment effect [6,9]. The index vessel is the vessel (either aorta, left

carotid, or right carotid) with the highest average mean of max TBR at baseline. The MDS is defined as the 1.5 cm segment within the artery that demonstrates the highest FDG activity. The MDS TBRmeanmax is calculated as a mean of maximum TBR values derived from 3 contiguous axial segments.

Secondary outcome definitions:

Change in vascular FDG-uptake: In addition to examining the MDS TBRmeanmax, five secondary vascular imaging endpoints will also be examined in exploratory analyses: 1) the TBRmeanmax across the entire index vessel; 2) carotid MDS TBRmeanmax; 3) aortic MDS TBRmeanmax, 4) TBRmeanmax across the entire carotid vessel (R and L averaged); and 5) TBRmeanmax across the entire aorta. These provide complementary information to the primary outcome. As a sensitivity analysis, we will also examine the target measures alone for the MDS of the index vessel, the entire index vessel, the carotid MDS, the aortic MDS, the entire carotid, and the entire aorta.

Change in DAS28: The DAS28 is calculated as noted below. The TJC and SJC were assessed by a blinded metrologist, and the patient rated their PtG based on a 0-100 VAS scale. The CRP (mg/dl) will be estimated in the central lab at the end of the study. The change in DAS28 will be calculated as the measure at the time of the final FDG PET/CT scan minus the baseline measure:

$$\text{DAS-28-CRP} = (0.56 \times \text{TJC28}) + (0.28 \times \text{SJC28}) + (0.36 \times \ln(\text{CRP}+1)) + (0.014 \times \text{PtG}) + 0.96$$

In addition to analyzing the full DAS28 score, each component of the DAS28 will be analyzed separately.

Change in biomarker levels: We will primarily test the Vectra-DA platform as it has been found to be a valid correlate of disease activity in RA[11] and is already used in clinical rheumatology practice. The Vectra-DA consists of 12 individual markers. All have been noted to also be markers of CVD risk[12]. Testing procedures for the Vectra-DA platform have been well described in the literature[13]. These assays will be run by Crescendo/ Myriad. The Vectra-DA will be used in Aim 2b to categorize disease activity as follows: remission < 25 or low 26-29 versus moderate 30-44 or high > 44 according to the manufacturer's algorithm[11,14]. However, this algorithm may not be optimal for assessing relationships to vascular inflammation. Since Cresendo/Myriad will provide us with the data for each individual protein at each measured timepoint, in a series of exploratory analyses, we will analyze each of the 12 components alone as a continuous variable and in various combinations to determine whether a different algorithm of the Vectra-DA components may be a better correlate of vascular inflammation.

Given the limited number of proteins in Vectra-DA, examination of a broader array of proteins involved in inflammation, endothelial activation, lipid processing, immunological pathways, etc, is desirable to enable the development of a novel multi-marker of vascular inflammation in RA. We will also examine a broader multi-analyte panel, the RBM Discovery Map. This will include targeted study of candidate markers based on systematic literature review, as well as a secondary analysis using a biomarker discovery approach.

Change in articular FDG uptake: While methods for analyzing articular FDG uptake are just emerging, there has been a validated methodology published during the conduct of the TARGET trial. [15] We will use the method as described below. This measure includes the following steps:

1. View all images in three orthogonal planes (axial, coronal, and sagittal);
2. Determine if a joint has synovitis by examining for increased FDG uptake in the anatomic region of the synovium compared with normal regional tracer accumulation;
3. Overlay a volume of interest (VOI) for FDG positive joints over the joint synovium on the PET images;
4. Create an iso-contour VOI for FDG positive joints that included all voxels above 42% of maximum;
5. Calculate SUVmax using the following formula:
    a. $\text{SUVmax} = \frac{\text{maximum activity in the VOI (MBq/mL)}}{\text{Injected dose (MBq)/body weight (g)}}$
6. Calculate SUVmax including 68 joints, except DIP in the hands and mid-tarsal joints; and
7. Repeat this process independently by another nuclear medicine physician or the same physician separated in time.

Based on the above steps, several measures will be calculated:

a. PET28: # of FDG positive joints out of 28 – primary articular endpoint
b. PET68: # of FDG positive joints out of 68
c. SUV28: mean SUVmax of 28 joints
d. SUV68: mean SUVmax of 68 joints

Thus, we will use the change in PET28 (final FDG PET/CT scan minus baseline FDG PET/CT) as the primary articular outcome (one of the secondary trial outcomes) and change in the other measures (b – d above) as secondary articular outcomes.

V. SAMPLE SIZE (Re-estimated)

In our original grant application, we proposed a sample size of 100 participants in each arm (200 total recruitment target) using information from previous data in RA patients which observed a baseline MDS TBRmax of 2.51 (SD of 0.33) and a 0.46 reduction after 8 weeks of a TNFi[16]. After discussions with the DSMB, we re-estimated sample size using blinded data from 45 subjects in the TARGET trial to estimate the SD and found **a standard deviation across scans of 0.291**. Thus, we re-estimated the required sample size during the trial using a t-test to compare MDS TBRmeanmax in the two arms at 6-months to generate a conservative power estimate with various sample sizes (see **Table 2**).

With an effective sample size of 126 (150 randomized minus 24 subjects (~16%) who we anticipated will drop-out or cross-over), the trial has a 99% power to detect an absolute difference of 0.17 between the 2 arms. This difference corresponds to the effect observed in the prior study by Maki-Petaja for TNFi's[16]. This difference between the arms would be a clinically important difference, on the same order for what was observed between a low-dose statin and a high-dose statin[17], a contrast with known clinical significance[18]. Power will be sufficient for the per protocol secondary analysis, including only the approximately 75% of subjects who remain compliant with study treatment. This pre-specified analysis will include an estimated 56 subjects per arm and will have 99% power to detect a difference of 0.17 in MDS TBRmeanmax. Since these are pre-specified secondary endpoints, p-values will not be adjusted.

Table 2: Power Calculation for Aim 1 (ANCOVA adjusting for baseline)

Between group difference	N=150	N=126	N=112
0.19	99%	99%	99%
0.18	99%	99%	99%
0.17	99%	99%	99%
0.16	99%	99%	98%
0.15	98%	98%	96%

VI. ANALYSIS PLAN

We will compare the baseline characteristics of those randomized to TNFi + MTX versus triple therapy using chi-squared tests (or Fisher's exact tests when cell counts were <5) for categorical variables and Wilcoxon rank sum tests for continuous variables. We will use the same methods to compare the baseline characteristics of individuals who completed the study protocol to those who did not.

We describe the analyses below with respect to the expected publications. While all of the aspects of the Specific Aims are included, they have been slightly re-organized and enhanced with more state-of-the-art statistical analyses (i.e. causal mediation analyses).

Analysis 1

This analysis will address the goals specified in Aim 1 and Aim 2a. Overall, this analysis will determine the effects on vascular inflammation of TNFi + MTX versus triple therapy (Part I) and explore whether these effects are mediated by changes in disease activity (Parts II and III). We will follow an approach similar to that used in the recent mediation analyses of the CANTOS trial anemia data[19,20].

Part I: The first part of this analysis will use an ANCOVA model estimating the change in index vessel MDS TBRmeanmax as a function of the baseline index vessel MDS TBRmeanmax, treatment group, and the randomization strata: 1) statin use at baseline, 2) oral steroid use at baseline, and 3) HCQ use at baseline. The null hypothesis is no association between treatment assignment and follow-up MDS TBRmeanmax. A p-value threshold of 0.05 for a two-sided test will be used to determine statistical significance. The primary analysis will only include participants with imaging data at baseline and follow-up. As a sensitivity analysis, we will exclude outliers defined as individuals with an Extreme Studentized Deviate (ESD) statistic for change in MDS TBRmaxman of greater than 3.46 [21]. We will also examine whether there were statistically significant changes in MDS TBRmeanmax within each arm from baseline to follow-up.

Not all participants will have a follow-up MDS TBRmeanmax because some participants will fail to complete the study protocol. We will perform a series of sensitivity analyses to explore how robust our results are to these missing data. First, we will impute the average MDS TBRmeanmax from all individuals with available follow-up MDS TBRmeanmax as the follow-up MDS TBRmaxmax for any individual missing this value. Second, we will perform multiple imputations using a Markov chain Monte Carlo technique[22] to impute follow-up MDS TBRmeanmax, including baseline MDS TBRmeanmax, randomized treatment assignment, and other baseline characteristics in our imputation model. Third, we will compare the baseline characteristics of individuals who complete the protocol and those who do not complete the protocol using t-tests and Fisher's exact tests as appropriate. If we observe any p-values <0.10, we will include that variable as a covariate in an ANCOVA model estimating the change in index vessel MDS TBRmeanmax as a function of the baseline index vessel MDS TBRmeanmax among individuals who completed the protocol.

Due to COVID-19, we amended our protocol in April 2020 and allowed individuals to receive their follow-up scan after up to 36 weeks of treatment. We will perform additional analyses controlling for length of time between PET/CT scans as well as analyses restricting to those who received their follow-up scan between 22-26 weeks as specified in the original protocol. Additional analyses will explore secondary vascular imaging endpoints. We will repeat the main analysis of the primary outcome in the subgroup compliant with study treatment (threshold of >80% of treatment days covered for all medications). Several secondary per protocol adherence analyses will also be pursued. If baseline imbalances are observed between the per protocol treatment groups, these variables will be included as covariates in our regression models We will also include analyses stratified by the following factors: achievement of LDAR; serologic status (seropositive versus seronegative); at least one CV risk factor (e.g., hypertension, diabetes, tobacco use, or hyperlipidemia); steroid use at baseline, gender, age (< median age versus ≥ median age), disease duration (< 5 years versus ≥ 5 years) and statin use at baseline (yes versus no). We will formally test for interaction between randomized treatment assignment and these factors by using an interaction term. An additional secondary analysis will compare the primary vascular inflammation outcome (MDS TBRmeanmax) change between adalimumab and etanercept users. These subgroup analyses will be underpowered and exploratory; we will report nominal p-values.

A sensitivity analysis for the impact of any missing covariate data will also be done using the same assigned treatment categories, but using all randomized participants with multiple imputations using a Markov chain Monte Carlo technique[22]. This will be most relevant for the subgroup analyses.

Part II: The next part of this analysis will address whether the magnitude of the treatment response achieved by individuals is related to their vascular inflammation response (baseline to final scan). This analysis will divide the participants into four groups according to randomized treatment assignment and whether they achieved low disease activity or remission (LDAR; DAS-28-CRP ≤ 3.2) versus remaining at moderate-high disease activity (MHDA; DAS-28-CRP > 3.2) as defined by the DAS-28-CRP at 18 weeks. The four groups will be defined as: (1) triple therapy remaining in MHDA (reference group); (2) triple therapy achieving LDAR; (3) TNFi + MTX remaining in MHDA; and (4) TNFi + MTX achieving LDAR. We chose to examine the DAS-28-CRP at 18 weeks as the mediator because we want to capture the effects of the treatment during the course of the study but also have the DAS-28-CRP measured prior to the assessment of the outcome. In the event that a participant drops out of the study prior to the 18-week assessment, we will use the DAS-28-CRP measure from the visit prior to the final imaging assessment.

We will use an ANCOVA model estimating the change in index vessel MDS TBRmeanmax as a function of the baseline index vessel MDS TBRmeanmax, randomization strata, the four groups combining information on treatment assignment and treatment response described above, a term for length of time between baseline and follow-up PET/CT scan, and the following potential confounders as measured at baseline: age, gender, disease duration, smoking status, serologic status, and body mass index.

Part III: This final part will use formal causal mediation analysis to control for baseline characteristics that might independently influence treatment response and to control for the effect of treatment. We see Part III as a complement to Part II, where Part II may be easiest to communicate to readers and Part III is statistically more sophisticated. Recent causal inference literature has highlighted potentially limitations to methods traditionally used to study mediation and have proposed new methodological approaches[23]. To implement these approaches, we will use Valeri and Vanderweele's SAS macro for causal mediation analysis. Because mediation analysis is sensitive to distributional assumptions, we will separately consider both a dichotomous and continuous measure of treatment response. First, we will use as a mediator an indicator of whether a randomly assigned participant achieves LDAR at 18 weeks (dichotomous mediator). Alternatively, as the DAS-28-CRP is a continuous measure of change in disease activity, we will calculate the natural log of the DAS-28-CRP at 18 weeks minus the DAS-28-CRP at baseline. For both measures of mediation, we will include a term for length of time between baseline and follow-up PET/CT scan and control for the following potential confounders as measured at baseline: age, gender, disease duration, smoking status, serologic status, prednisone use, statin use, and BMI.

Analysis 2

We will use a similar overall methodology as described in Analysis 1, but the outcome of interest will be change in joint inflammation as measured by articular FDG PET/CT (PET-28). While this analysis was not called out as a Specific Aim, it will be of great interest to the rheumatology community, because articular FDG PET/CT is of increasing interest as an RA trial outcome. First, we will perform the same analyses as described in Part I above which will examine the effect of randomized treatment assignment on change in articular inflammation. Next, we will perform the analyses described in Part II using either DAS-28-CRP or Vectra-DA measured at 18 weeks as our mediator. As described above, we will consider both a dichotomous and continuous measure of the DAS-28-CRP. For Vectra-DA, our dichotomous definition of the mediator will be Vectra-DA score less than or equal to 29 at 18 weeks versus a score of 30 or greater at 18 weeks. As a normalized, continuous measure of the mediator, we will calculate the natural log of Vectra-DA at 18 weeks minus Vectra-DA at baseline.

Analysis 3

While prior work has explored the association between Vectra-DA and DAS-28-CRP, less is known about the association between Vectra-DA and measures of vascular inflammation and articular inflammation. Our third set of analyses are exploratory and will leverage data from the TARGET trial to obtain additional insights into potential associations between biomarkers and vascular and articular inflammation.

These analyses will pool data across both treatment arms of the trial. We will compare several aspects of the Vectra-DA to several aspects of the MDS TBRmeanmax vascular assessment. The Vectra-DA consists of 12 analytes that are typically combined as a continuous score or put into three categories (low, moderate, and high). In addition, each of the 12 analytes that comprise the Vectra-DA can be considered as individual predictors.

Several analyses are planned. First, baseline (pre-treatment) measurements of the continuous Vectra-DA will be compared with the baseline vascular inflammation assessment. This cross-sectional relationship will be assessed using linear regression with vascular inflammation as the dependent variable and continuous Vectra-DA as the independent variable of interest. The following baseline variables will be considered as covariates:, age, sex, blood pressure, low-density lipoprotein cholesterol, diabetes (yes/no), RA disease duration, smoking status, serologic status, baseline disease activity (DAS-28-CRP), prednisone use, statin use (none, low, or moderate intensity), and BMI.

Second, changes in Vectra-DA over time (continuous measurement) will be compared with the 6-month change in MDS TBRmeanmax. We will use linear mixed models and include all Vectra-DA assessments as independent variables. The same set of covariates as noted above will be included as well as baseline MDS TBRmeanmax, treatment arm, and length of time between baseline and follow-up PET/CT scan.

In a series of exploratory analyses, we will also consider each of the 12 analytes in separate models. These models will include all of the above covariates.

The above analyses will be repeated using the measures of articular inflammation (PET-28) as the outcomes instead of MDS TBRmeanmax.

Analysis 4
These analyses are exploratory and were not described as Specific Aims. We will compare change in vascular inflammation (MDS TBRmeanmax) with change in articular inflammation (PET-28). This will be examined across both treatment arms and then individually, by treatment arm. The change in articular inflammation (PET-28) will be considered the independent variable and the change in MDS TBRmeanmax will be the dependent variable. These analyses will use linear regression and will include the same set of covariates as noted in Analysis 3. We will also examine subgroups, including by prednisone use (yes/no), statin use (yes/no), treatment arm (triple therapy/TNF antagonist), serologic status, and sex. We will also explore whether there is a treatment interaction by including an interaction term between treatment and articular inflammation.

VII. SAFETY

Since we tested FDA approved medications for their approved indications, there were no safety concerns to be addressed by an interim analysis and possible trial cessation. There are strong arguments against stopping small trials because of interim analyses[24,25]. The DSMB did not choose to impose stopping rules. Two sample tests of proportion will compare safety data across treatment groups, including infection and cancer events. We monitored the distribution of baseline variables across treatment groups and compliance with study protocol. These analyses were performed by the DCC and provided to the DSMB.

References

1. 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 and rheumatism.* 1995;38(1):44-48.
2. Centola M, Cavet G, Shen Y, et al. Development of a multi-biomarker disease activity test for rheumatoid arthritis. *PloS one.* 2013;8(4):e60635.
3. Fransen J, van Riel PL. The Disease Activity Score and the EULAR response criteria. *Rheumatic diseases clinics of North America.* 2009;35(4):745-757, vii-viii.
4. Marincheva-Savcheva G, Subramanian S, Qadir S, et al. Imaging of the aortic valve using fluorodeoxyglucose positron emission tomography increased valvular fluorodeoxyglucose uptake in aortic stenosis. *J Am Coll Cardiol.* 2011;57(25):2507-2515.
5. Fifer KM, Qadir S, Subramanian S, et al. Positron emission tomography measurement of periodontal (18)f-fluorodeoxyglucose uptake is associated with histologically determined carotid plaque inflammation. *J Am Coll Cardiol.* 2011;57(8):971-976.
6. Fayad ZA, Mani V, Woodward M, et al. Safety and efficacy of dalcetrapib on atherosclerotic disease using novel non-invasive multimodality imaging (dal-PLAQUE): a randomised clinical trial. *The Lancet.* 2011.
7. Rogers IS, Nasir K, Figueroa AL, et al. Feasibility of FDG imaging of the coronary arteries: comparison between acute coronary syndrome and stable angina. *JACC. Cardiovascular imaging.* 2010;3(4):388-397.
8. Tawakol A, Migrino RQ, Bashian GG, et al. In vivo 18F-fluorodeoxyglucose positron emission tomography imaging provides a noninvasive measure of carotid plaque inflammation in patients. *J Am Coll Cardiol.* 2006;48(9):1818-1824.
9. Tawakol A, Fayad ZA, Mogg R, et al. Intensification of Statin Therapy Results in a Rapid Reduction in Atherosclerotic InflammationResults of a Multicenter Fluorodeoxyglucose-Positron Emission Tomography/Computed Tomography Feasibility Study. *J Am Coll Cardiol.* 2013;62(10):909-917.
10. Rudd JH, Narula J, Strauss HW, et al. Imaging atherosclerotic plaque inflammation by fluorodeoxyglucose with positron emission tomography: ready for prime time? *J Am Coll Cardiol.* 2010;55(23):2527-2535.
11. Curtis JR, van der Helm-van Mil AH, Knevel R, et al. Validation of a novel multibiomarker test to assess rheumatoid arthritis disease activity. *Arthritis care & research.* 2012;64(12):1794-1803.
12. van Holten TC, Waanders LF, de Groot PG, et al. Circulating biomarkers for predicting cardiovascular disease risk; a systematic review and comprehensive overview of meta-analyses. *PloS one.* 2013;8(4):e62080.
13. Costanzo P, Cleland JG, Vassallo E, Perrone-Filardi P. Questioning the predictive role of carotid intima-media thickness. *Future cardiology.* 2011;7(5):651-656.
14. van der Helm-van Mil AH, Knevel R, Cavet G, Huizinga TW, Haney DJ. An evaluation of molecular and clinical remission in rheumatoid arthritis by assessing radiographic progression. *Rheumatology (Oxford).* 2013;52(5):839-846.
15. Lee SJ, Jeong JH, Lee CH, et al. Development and Validation of an (18) F-Fluorodeoxyglucose-Positron Emission Tomography With Computed Tomography-Based Tool for the Evaluation of Joint Counts and Disease Activity in Patients With Rheumatoid Arthritis. *Arthritis & rheumatology.* 2019;71(8):1232-1240.
16. Maki-Petaja KM, Elkhawad M, Cheriyan J, et al. Anti-tumor necrosis factor-alpha therapy reduces aortic inflammation and stiffness in patients with rheumatoid arthritis. *Circulation.* 2012;126(21):2473-2480.

17. Tawakol A, Fayad ZA, Mogg R, et al. Intensification of statin therapy results in a rapid reduction in atherosclerotic inflammation: results of a multicenter fluorodeoxyglucose-positron emission tomography/computed tomography feasibility study. *J Am Coll Cardiol.* 2013;62(10):909-917.
18. Ribeiro RA, Ziegelmann PK, Duncan BB, et al. Impact of statin dose on major cardiovascular events: a mixed treatment comparison meta-analysis involving more than 175,000 patients. *International journal of cardiology.* 2013;166(2):431-439.
19. Vallurupalli M, MacFadyen JG, Glynn RJ, et al. Effects of Interleukin-1beta Inhibition on Incident Anemia: Exploratory Analyses From a Randomized Trial. *Ann Intern Med.* 2020;172(8):523-532.
20. VanderWeele TJ. *Explanation in causal inference : methods for mediation and interaction.* 2015.
21. Rosner B. *Fundamentals of Biostatistics, 8th edition.* Boston, MA: Cengage Learning; 2016.
22. Horton NJ, Kleinman KP. Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *The American statistician.* 2007;61(1):79-90.
23. VanderWeele T. *Explanations in Causal Inference* Oxford University Press 2015.
24. Pocock SJ. When to stop a clinical trial. *BMJ (Clinical research ed.* 1992;305(6847):235-240.
25. Green SJ, Fleming TR, Emerson S. Effects on overviews of early stopping rules for clinical trials. *Statistics in medicine.* 1987;6(3):361-369.