**SUPPLEMENTARY FILE**

**Additional information to Methods**

*Study medication Rituximab*

Rituximab (RTX, MabThera®/Rituxan®) is a chimeric immunoglobulin G1 kappa (IgG1κ) monoclonal antibody directed against the human CD20 antigen, comprising murine light and heavy chain variable regions (14%) and human G1 heavy chain and κ light chain constant regions (86%). Rituximab binds specifically to human CD20, expressed on the majority of B cells, from the pre–B cell stage through differentiation into mature or memory B cells, but is not expressed on plasma cells. Treatment with rituximab induces a rapid and sustained depletion of peripheral CD20+ B cells by several potential mechanisms, including complement-dependent cytotoxicity (CDC), antibody-dependent cellular cytotoxicity (ADCC), and apoptosis.

*Study procedures*

*Study visits*

In addition to the screening and baseline visit, follow up visits were planned at 4, 8, 12, 16, 24, 36 and 52 weeks. After that, subjects were checked on a yearly basis. In addition to the planned follow up visits, subjects were asked to contact members of the study team immediately in the case of a suspected arthritis. The subjects were then checked to verify the suspected arthritis within 2-3 days using the approach mentioned in the main manuscript.

*Joint count*

The 68 joints assessed included: 2 temporomandibular, 2 sternoclavicular, 2 acromioclavicular joints, 2 shoulders, 2 elbows, 2 wrists, 10 metacarpophalangeal joints of the hands, 10 proximal interphalangeal of the hands, the 8 distal interphalangeal joints of the hands, 2 hips, 2 knees, 2 ankles, 2 mid-tarsal joints, 10 metatarsophalangeal joints of the feet and 10 proximal interphalangeal joints of the feet.

*Detection of fine specificity of autoantibodies against citrullinated peptides*

Additional serum antibodies against the following citrullinated peptides and arginine-containing peptides were determined at baseline and follow up visits according to the methods mentioned in the main manuscript: anti-Fillaggrin 307-324, anti-Vimentin 60-75; anti-Fibrinogen β 36-52; anti-Fibrinogen β 573; anti-Fibrinogen β 591.

*Determination of IgM, IgA, and IgG rheumatoid factor by ELISA*

For determination of the presence and levels of IgM rheumatoid factor (IgM RF), the IgM RF ELISA kit (Hycor cat.no. FGA05/DIB002) was used according to instructions of the manufacturer. The cut off value was adjusted to 50 IU/ml instead of the suggested 25 IU/ml in order to increase the specificity. For IgA/IgG RF determination, the QUANTA Lite RF IgA (Inova 708695) and QUANTA Lite RF IgG (Inova 708685) ELISA tests were used.

*Determination of B lymphocytes and T lymphocytes by flow cytometry*

T lymphocytes (CD3+) and B lymphocytes (CD19+) were determined by flow cytometry using BD Multitest™ 6-color TBNK reagents with BD Trucount™ tubes (BD Biosciences cat. nr. 337166) according to the manufacturer’s instructions.

*Analysis B lymphocyte subtype frequencies*

At baseline, 6 months, 12 months, and at the moment arthritis developed, blood samples were collected to determine the B cell and T cell frequencies. In a subset of subjects (n=75) the expression of 28 markers was determined using Fluorescence-activated cell sorting (FACS) analysis on isolated peripheral blood mononuclear cells (PBMCs, isolated according to a standard density gradient protocol and cryopreserved until further use). B cells were defined as the CD19+CD20+ population. After washing the cells with PBS 0·5 % and BSA 0·01% NaN3, cells were stained for 30 minutes at room temperature and directly labelled with antibodies to stain for B cells markers: IgM FITC , IgD PE, CD20 Cy5.5 PerCP, CD38 Pe-Cy7, CD19 Alexa 700 (BD Biosciences, Breda, the Netherlands) and the activation status of B cells : CD19 FITC, CD40 PE (eBioscience), CD25 APC CD69 PerCP (BD), CD80 Pe-Cy7, CD5 APC-H7, HLA-DR Alexa 700 (eBioscience). After incubation, the cells were washed and analysed using a FACS CANTO II (BD Biosciences, Breda, the Netherlands). Data were analysed using FlowJo software (Treestar Inc., Ashland, OR). Complete depletion of B cells was defined as a B cell count of < 0·0001x109/L.

***Statistical analysis***

*Power analysis to determine the number of subjects*

Time-dependent hazard info for log-rank survival test

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  End of period, time t (years) | 0·0 | 1·0 | 2·0 | 3·0 | 4·0 |
|  Accrual (% of total) | 0 | 75 | 25 | 0 | 0 |
|  Group 1 exponential hazard rate, λ1 | 0·0000 | 0·0513 | 0·1231 | 0·1542 | 0·1823 |
|  Group 2 exponential hazard rate, λ2 | 0·0000 | 0·0202 | 0·0206 | 0·0317 | 0·0328 |
|  Group 1 expected % surviving time t | 100·0 | 95·0 | 84·0 | 72·0 | 60·0 |
|  Group 2 expected % surviving time t | 100·0 | 98·0 | 96·0 | 93·0 | 90·0 |
|  Common exponential dropout rate, d | 0 | 0 | 0 | 0 | 0 |

***Supplementary table 1*** **Determination of sample size using the log-rank test for equality of survival curves with a p=0·050 two-sided significance level when an estimated follow period of 4 years is applied**

***Additional information to Results***



*B cell subset analyses*

A more detailed analysis of subsets of B-cell populations could be performed in 45 patients (n=18 rituximab, n=27 placebo) using FACS analysis, based on availability of samples. There was a statistically significant decrease in the frequencies of CD19positive B cells, most pronounced after 6 months, returning to near baseline levels at 12 months after treatment in the rituximab group*.* The percentages of naïve and unswitched memory B cells decreased significantly. The plasmablast subsets (CD38 positive), regulatory B-cell containing fractions (defined by CD19positiveCD27positive and CD19positiveCD5positive populations) and the switched (IgD negative) B cell population, both naïve and memory, were significantly increased 6 months after rituximab treatment when compared to the placebo treated individuals. The frequencies returned to normal at 12 months after treatment, alongside an observed repopulation of B cell numbers *Fig. S1 A-G)*. There was no significant change in the percentages of B and T-cell subsets in subjects treated with placebo, nor were the numbers and percentages of T cells in both treatment groups. In subjects treated with rituximab who developed arthritis during the follow up period, the CD27positiveIgDnegative (switched memory cell) population showed a significant decrease before the diagnosis of arthritis was established (p< 0·016), whereas the CD27negativeIgDpositive naive B cell population tended to be increased at the same time point (p=0·11). These changes were not observed in the individuals who developed arthritis and in the placebo group *(Fig. S2 A-G).*

**Figure Legends supplemental file**

**Figure S1 A-G Changes in B cell population subsets over time** FACS analysis of B-cell population subsets in 45 individuals (n=18 rituximab, n=27 placebo) at baseline, 6 months and 12 months after study treatment (rituximab (green) and placebo (red)). The thin lines represent frequencies measured in individuals, the thick lines represent the mean frequencies for each treatment. Vertical lines represent the 95% confidence intervals (CI). Percentages of CD19 positive CD20positive B cells (**A**), naïve and unswitched memory B cells (**B** and **C**, respectively), plasmablast subsets (CD38 positive; **D**), regulatory B-cells (CD19 positive CD27 positive, CD19 positive CD5 positive, **E** and **F**, respectively) and the switched (IgD negative) B cell population (**G**).

**Figure S2 A-G Changes B cell population subsets in relationship with arthritis development**

 B-cell population subsets in 45 individuals (n=18 rituximab, n=27 placebo) at baseline, 6 months and 12 months after study treatment (rituximab (green) and placebo (red)). The thin lines represent frequency measured in individuals, the thick lines represent the mean frequencies for each treatment. Vertical lines represent the 95% confidence intervals (CI). In the left panel, percentages of B cell subsets are shown in individuals who did not develop arthritis (x-axes represents months since inclusion), in the right panel results of percentages in individuals who developed arthritis are depicted (x-axes represents month prior to arthritis development).Percentages of CD19positive B cells (**A**), naïve and unswitched memory B cells (**B** and **C**, respectively)., plasmablast subsets (CD38 positive; **D**), regulatory B-cells (CD19 positive CD27 positive, CD19 positive CD5 positive, **E** and **F**, respectively) and the switched (IgD negative) B cell population (**G**).