Research Protocol

1. Project title: Evaluating treatment response mechanisms of Rituximab and Tocilizumab in Rheumatoid Arthritis in the R4RA randomised clinical trial

2. Summary of Research (Abstract)

Background
The proposed research capitalises on the completed NIHR-EME funded R4RA Randomised Clinical Trial (RCT) to evaluate the mechanisms of action of Rituximab (RTX) and Tocilizumab (TOC). R4RA [1] is the first synovial biopsy-driven multicentre RCT aimed at testing the hypothesis that specific cellular/molecular signatures in rheumatoid arthritis (RA) joint tissue (high v low B cells; high v low IL-6 signalling system) influence therapeutic outcome to RTX or TOC. Analysis of the R4RA trial has shown that in RA patients expressing low B cells / B-cell molecular signatures in synovial biopsy (the target for RTX) this drug was less effective compared to TOC, implicating that target expression levels in the disease tissue are involved in response/non-response to RTX and TOC. Here we propose an investigation to elucidate the mechanisms.

Research question
We will address 3 research questions testing the hypothesis that:
i) Low expression of RTX and TOC targets (CD20 and IL-6/IL6-R pathways) and/or critical downstream cells and signalling pathways (plasmablasts, plasma cells, IL-6R/gp130 complex) in RA synovium is associated with failure to respond to RTX and TOC;
ii) Failure to deplete synovial B cell clonotypes, plasmablasts, plasma cells (CD20-ve) or repopulation of pathogenic B cell clones from synovial niche or primary/secondary lymphoid organs is associated with persistent disease and treatment failure in RA;
iii) Blood markers of synovial pathotypes may contribute to enhance mechanistic understanding of response and help predict therapeutic outcome.

Aims and objectives
1. Elucidate mechanisms of response for TOC and RTX in RA through deep cellular and molecular phenotyping.
2. Determine whether early repopulation of specific B cell subsets (CD27+ memory B cells/plasmablasts) drive non-response to RTX, and whether contrasting B cell subsets mediate treatment failure to TOC.
3. Determine relationship between synovial vs blood plasmablast/plasma cell clonotype and response to RTX and/or TOC.
   a. Does the presence of long-lived plasma cell clonotypes in synovium mediate drug resistance?
   b. Does early repopulation of pathogenic clonotypes post-RTX or persistence of the same clonotypes post-TOC mediate drug resistance?
4. Define whether low expression of IL-6/IL-6R gene module and/or B-cell signatures predict poor response to TOC.
5. Test the hypothesis whether previously reported biomarkers in the peripheral blood (PB) act as surrogates of tissue pathology (akin to a 'liquid biopsy'), and are mechanistically linked to response to RTX and TOC.

Methods
1. Deep immunophenotyping synovial biopsies
Deep immune-histological characterization by multiplex IHC and Hyperion mass cytometry using antibody markers for B/T cell, plasma cell, macrophage, FDC (Follicular Dendritic cells) & fibroblast populations.
2. Correlate B cell differentiation states in blood and synovium with treatment outcome
Plasmablasts/ CD27+ memory B cell repopulation will be measured by high sensitivity flow cytometry (hsFACS) and compared with data from Multiplex IHC and Hyperion analysis of B cell subsets in synovium.

3. Correlate clonotype analysis in synovium and blood
B cell receptor repertoire sequencing and family tree analysis pre and post-treatment will be integrated with Multiplex IHC and Hyperion analysis of synovial B cell subsets and hsFACS of peripheral blood B cells to map B cell clonality in tissue and peripheral blood as a mechanism of non-response to RTX and TOC.
4. Bioinformatic Analyses
i) Bioinformatic synovial RNA-seq expression analysis for drug targets & linked signalling apparatus and response to TOC/RTX in R4RA.
ii) RTX predictive model and FANTOM5 B cell module will be validated in the STRAP study comparing RTX and TOC cohorts with further comparison against anti-TNF cohort. Optimal cut-off for both predictive models will be compared in R4RA and STRAP, and subject to meta-analysis.
5. Peripheral blood biomarker analyses
i) Luminex analysis of serum markers associated with synovial pathology, therapeutic response and up/downstream factors associated with B cell maturation and IL-6 secretion.
ii) q-PCR analysis of IFN response signature and B cell maturation associated blood transcripts pre/post treatment and with response.

Timelines for delivery
The project will be delivered over a 2 year period adhering to milestones and Gantt chart as documented (see main application).

Anticipated impact and dissemination
The project will improve scientific knowledge of the mechanisms of effective treatment in RA, underpinning future research funding and industry participation. Elucidating the mechanisms driving response/non-response to RTX and TOC could pave the way for identifying biomarkers for optimal stratification of patients. Results will be disseminated through the R4RA website, open access publications, presentations at national and international meetings, patient and public involvement (PPI) groups and through specialist rheumatology societies and national patient organisations.

3. Background and Rationale

Prevalence & size of problem
Rheumatoid arthritis (RA) is one of the most important chronic inflammatory disorders, affecting ~1% of adults (~500K in UK), leading to increased morbidity and mortality [2,3]. Although conventional synthetic (cs) and biologic (b) disease modifying anti-rheumatic drugs (DMARDs) have revolutionized the treatment of RA, ~40% are non-responsive to therapy [3,4]. This proportion goes up to 50% in patients who did not respond to csDMARDs and the first commonly used bDMARD: anti-TNF. Consequently, approximately one-third of patients are affected by disability, with work & social security costs being estimated at £3.8-4.75 billion/year, while direct medical costs exceed £800 million/year (NICE guideline NG100).

Knowledge gap & unmet medical need
As the mechanism of response/non-response in RA to DMARDs are still unknown and we lack predictive biomarkers of treatment response to individual drugs [5], current practice is based on a “trial and error” use of DMARDs leading to:

- unnecessary exposure to potentially toxic drugs with a low probability of response
- delay in disease control, progressive joint destruction and secondary disability linked to
- large waste of valuable NHS/societal resources.

Therefore, understanding the mechanisms of action of targeted bDMARDs remains a fundamental healthcare goal in order to identify predictive biomarkers able to inform therapy allocation and enrich likelihood of response.

To achieve this goal, since extensive research in peripheral blood (PB) biomarkers to predict treatment response to specific drugs has not been successful so far [5], we developed an integrated biopsy-driven randomised clinical trial (RCT) programme based on the hypothesis that

the expression levels of specific pathways targeted by bDMARDs at the main site of disease: the synovial tissue, would play a major role in drug response [6]. We pioneered the use of ultrasound-(US)-guided biopsy as a safe, well-tolerated method [7,8] that underpinned the first synovial biopsy-driven multicentre RCT worldwide funded by NIHR described below.

R4RA randomised clinical trial

The proposed research capitalises on the NIHR-EME funded R4RA Randomised Clinical Trial (EME ref 11/100/76): A Randomised, open labelled studying anti-TNFa inadequate responders to investigate the mechanisms for Response - Resistance to Rituximab (RTX) versus Tocilizumab (TOC) in RA (R4-RA)" (http://www.r4ra-nihr.whri.qmul.ac.uk/). R4RA is the first synovial biopsy-driven multicentre RCT worldwide aimed at testing the hypothesis that specific cellular and molecular signatures in the disease joint tissue (e.g. high v low/absent B-cells and associated signatures) critically influence therapeutic outcome to the above commonly used biologic (b)-DMARDs [1].

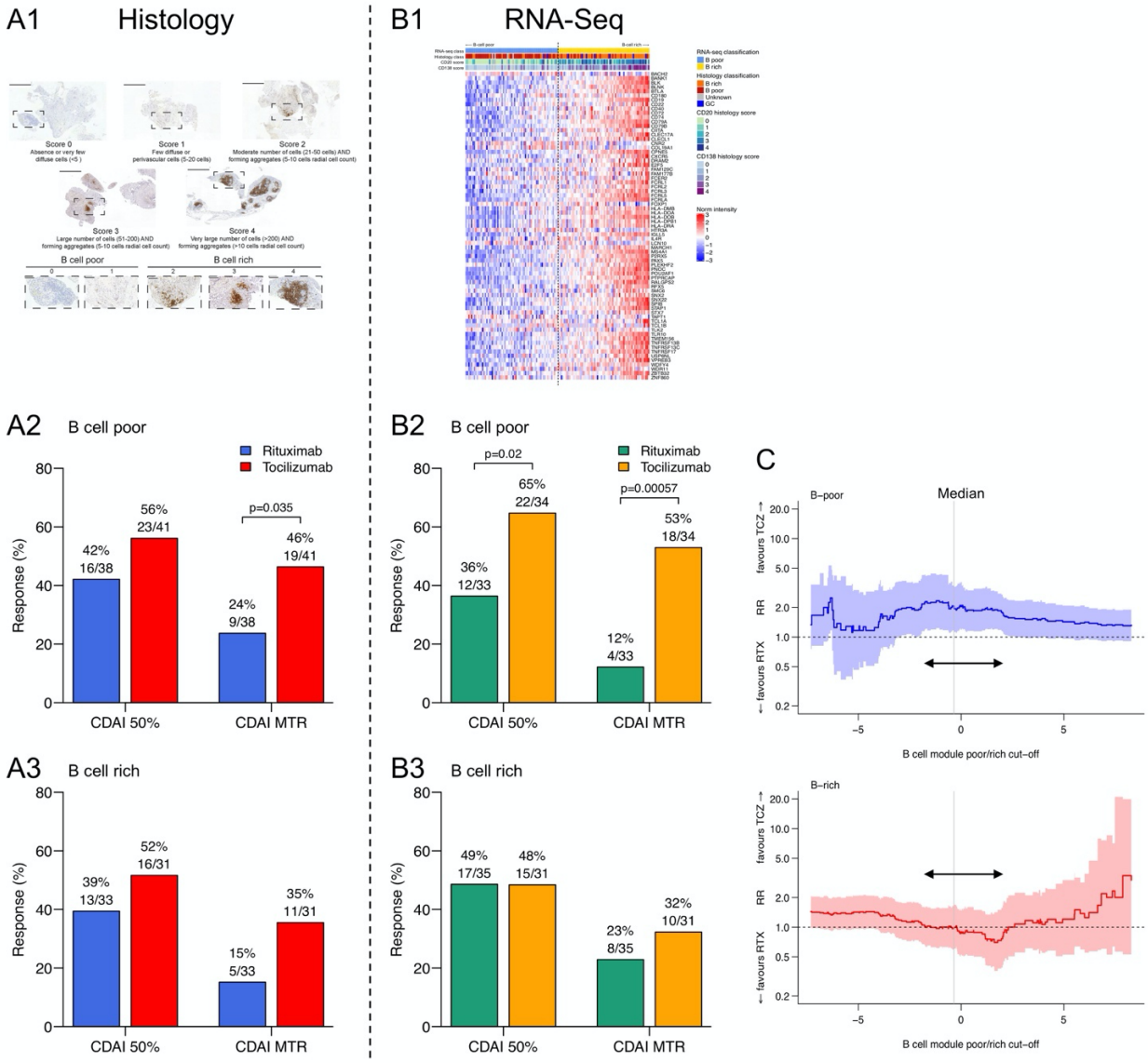


Fig 1. (A1) Histology atlas showing semi-quantitative 0-4 grading of B cells in synovial biopsies from RA patients in the R4RA trial. (A2,3) Response rates determined as CDAI ≥ 50% improvement or CDAI major treatment response (MTR) in patients randomised to Rituximab (RTX) or Tocilizumab (TOC), stratified by synovial histology into (A2) B cell poor or (A3) B cell rich groups. (B1) Heatmap showing FANTOM5 B cell module genes for patients in R4RA classified into B cell poor/rich by RNA-Seq. (B2) Response rates for CDAI ≥ 50% improvement or CDAI MTR in patients randomised to RTX or TOC and stratified into (B2) B cell poor or (B3) B cell rich by RNA-Seq B cell gene module. (C) Plot showing how the risk ratio for TOC vs RTX for CDAI 50% responders in B cell poor (blue) and B cell rich (red) groups varies if the RNA-seq B cell module cut-off is varied (x axis). Vertical grey line shows median RNA-seq B module score corresponding to (B2) original bar plots of study using median B cell score as cut-off.

The study was powered to test superiority of TOC over RTX in the B-cell-poor population at 16 weeks. Primary efficacy analysis evaluated the number of patients meeting primary end-point: CDAI≥50% improvement from baseline. Patients could also be deemed non-responders, as pre-defined in the Protocol, if they achieved CDAI improvement ≥ 50% from baseline but did not reach low disease activity CDAI<10.1, defined as CDAI major-treatment-response (CDAI-MTR).

Analysis of the R4RA trial [1] has shown that in RA patients with low B cells in synovial biopsy classified by histopathology (number of CD20+ve B cells), RTX was not significantly different compared to TOC at achieving ≥ 50% improvement in Clinical Disease Activity Index (CDAI) (42% v 56%). However, RTX was significantly less effective compared to TOC for the proportion of patients achieving more stringent low disease activity i.e. CDAI-MTR: 24% v 46% (p=0.035 (Fig. 1A). This result was even more striking when an RNA-Seq based B-cell poor/rich classification as pre-defined in the Protocol was utilised: with both endpoints reaching statistical significance (CDAI)≥50% (36% v 65%, p=0.020) and CDAI-MTR (12% v 53%, p=0.0005) (Fig 1B). Logistic regression analysis showed a statistically significant interaction between RNA-seq defined B-cell subgroups and randomised drug (p=0.049) was observed when using CDAI-MTR (p=0.049), providing evidence that the difference between RTX and TOC was statistically different between molecularly defined RNA-seq B-cell-poor and B-cell-rich stratified groups.

The study was not powered to evaluate comparative efficacy of either drug in the B-cell-rich cohort, however, assessment of CDAI≥50% response and CDAI-MTR at 16 weeks was also carried out as a supplementary analysis for completeness, where the response rate of RTX was compared to TOC, showing no significant difference between the drugs in this group.

Similar findings were observed when using continuous response outcomes showing that greater improvement in CDAI, DAS28-ESR and DAS28-CRP was observed in TOC treated patients in the RNA-Seq classified B cell poor group (Table 1 below).

Continuous secondary endpoints at week 16	Rituximab (n=33)	Tocilizumab (n=32)	Least squares mean difference (95% CIs)	Unadjusted p value
CDAI, least squares mean change	-10·85 (1·98)	-17·17 (2·01)	6·32 (0·67 to 11·97)	0·029
DAS28(ESR), least squares mean change	-1·26 (0·23)	-2·8 (0·23)	1·54 (0·89 to 2·18)	0·000012
DAS28(CRP), least squares mean change	-1·13 (0·22)	-2·14 (0·22)	1·01 (0·38 to 1·64)	0·0021

While pilot data was used prior to the start of R4RA clinical trial to determine the cut-off for histological classification of B cell poor by IHC (Score 0-1 Fig.1A1) and rich (Score 2-3-4 Fig.1A1), no “gold standard” existed in the literature, nor data was available to determine an optimal cut-off for RNA-Seq B cell module. Since RNA-Sequencing was performed after the end of the R4RA study, the median RNA-Seq gene module value was used to delineate B-cell poor and B-cell rich subgroups, to avoid selection bias. Further post-hoc statistical analysis to determine the effect of varying the RNA-Seq B cell module cut-off across its range (Fig. 1C) shows that the median value is approximately optimal, corresponding to the middle of a plateau in the relative risk between the 2 drugs at this point. Thus, varying the RNA-Seq B cell poor/rich cut-off over a 20% range of values did not significantly alter the outcome of the study based on the primary outcome measure.

Histology analysis of paired week 16 synovial biopsy was available in 41 patients treated with RTX and 24 patients treated with TOC. In the RTX cohort, when patients were stratified into week 16 responder (n=15) vs non-responder (n=26) (CDAI≥50% improvement) groups CD20+ and CD79a+ve B cells decreased significantly between baseline and 16 weeks in both groups, although the percentage decrease was larger in the responder group (CD20+ -96% vs -72%, and CD79a -56% vs -43%, p<0.001).

Overall these results suggest that additional mechanisms linked to the biomarkers picked up by the 73-gene molecular module associated with B cell differentiation/lineage are involved in response/non-response to RTX and TOC. Histology analysis of paired biopsies suggests that deeper depletion of B cells in synovial tissue is associated with enhanced response to RTX.

The results can also be interpreted that, in line with the predicted biological effects of blocking IL-6 signalling [9-11], TOC affects both B-cell and non-B-cell pathways. Thus, more complex molecular circuits are involved in the mechanisms of response. Understanding the mechanisms of response/non-response for both RTX and TOC is paramount in order to identify clinically useful biomarkers and inform our understanding of RA pathogenesis to advance future therapeutics.

Capitalising on the wealth of longitudinal clinical trial data (fully recruited to target) and the R4RA associated comprehensive biomedical resource, which includes 1431 sequential blood samples over 48 weeks follow-up with matched synovial biopsies (237 in total: 164 at baseline, 65 at primary end-point, 8 at 48 weeks), there is an opportunity to exploit this unique platform to test hypotheses and further enhance our understanding of the mechanisms underlying response/non-response to RTX and TOC.

We will validate specific findings from R4RA including the RNA-Seq B cell classification and RTX 35-gene predictive model in the Stratification of Biologic Therapies for RA by Pathobiology (STRAP) study funded by the MRC. STRAP is a randomised biopsy-driven stratification trial in DMARD-inadequate responder RA patients randomised to Etanercept, TOC or RTX. Thus, STRAP has response data in patients treated with TOC or RTX in which to validate mechanistic findings and response biomarkers derived from R4RA. Recruitment for STRAP is already complete and last remaining patients are undergoing follow-up. Clinical database lock is estimated to occur in autumn 2020. RNA-Sequencing (also funded by MRC, added value) of synovial biopsies from STRAP is already in progress and data are awaited in the next few weeks.

Systematic literature review and Supporting Evidence

To inform our research plan we first carried out a systematic literature review under the aegis of an international project entitled “The precision medicine in profiling the patients regarding therapy. From Pathology to treatment, what evidence in rheumatic and autoimmune diseases?”. Strict methodology based on Delphi Technique was followed, starting from the results of a systematic review of available literature, with the aim of producing a set of statements summarising the consensus among the Experts, as previously reported [12]. Within the RA working group, the following research questions were identified: i. “In RA patients, does B cell analysis help to predict treatment response to B cell depletion with Rituximab?” and ii. “Do inflammatory markers predict the response to IL-6 inhibitors?”. Search strings were developed using a PICO strategy and the following criteria were used: a. Databases: Medline via Pubmed, Web of Science, Cochrane database; b. Time-frame: July 1, 2009 – July 1, 2019; c. Study design: RCT, Observational studies or case series; d. Language: English language; e. Population: RA patients.

Narrative reviews, editorials, scientific conference abstracts, case reports and preclinical studies were excluded. Following initial screening of titles and abstracts, full-text were analysed to confirm eligibility. Any uncertainties and/or disagreements were resolved by discussion until reaching a final consensus. Data extraction was then performed and independently verified. The hierarchy of study types was indicated by levels of evidence suggested by Oxford University [13]. Following presentation of the results, statements were discussed, eventually reformulated, and voted through a Delphi-method reaching a consensus involving a panel of International Experts (unpublished data Rivellese, Pitzalis *et al*, manuscript in preparation).

In RA, RTX induces a depletion of circulating B cells in all treated patients, thus the level of B cell depletion in the peripheral blood after the first infusion does not predict response (2b). High sensitivity flow cytometry and the pre-treatment analysis of specific B cell subpopulations or B cell lineage markers help to identify responders (level of evidence 3b). Upon treatment with RTX, a variable depletion of synovial B cells has been described and the presence of specific B cell sub-populations in synovia, such as CD79a+ B cells and pre-plasma cells, has been associated with treatment response (level of evidence 4). There is no association between inflammatory markers at baseline (CRP, ESR) and attainment of response when using composite measures that include CRP and/or ESR e.g. DAS28 remission or EULAR response, following TOC treatment in RA (level of evidence 2b). Higher levels of CRP are associated with lower TOC discontinuation in RA (level of evidence 4).

RTX is an established treatment for RA, as shown by several randomized clinical trials that confirmed its efficacy. RTX targets CD20+ B cells, inducing their depletion in the peripheral blood. However, pharmacodynamics and post-hoc analyses of these trials have shown that the levels of pre-treatment B cells or the depth of their depletion, as measured by conventional flow cytometry, do not associate with treatment response [14-22]. The use of high sensitivity flow cytometry might help to set a stricter cut-off level for B cell depletion, however this has only been tested in small observational studies with somehow contradictory results [23-25]. Higher levels of specific B cell sub-populations in the peripheral blood, such as CD27+ memory B cells [26,27] and pre-plasma cells [28] are associated with reduced response to RTX. Additionally, a number of circulating markers have been linked to the response to RTX, such as mRNA levels of IgJ [29], polymorphism of BAFF [30], Fcgamma R IIIa [31] and IL-6 [32], whole blood transcriptomic signatures, serum IL-33 [33] and CCL19 levels [34]. Importantly, most of these observations originate from individual studies with relatively small numbers, and have not been reproduced or validated, therefore their clinical utility for patient stratification to RTX is limited.

In contrast to the complete depletion of peripheral blood B cells, a variable depletion of B cells and other immune cells in synovia following treatment with RTX has been described [35-42]. In particular, the number of pre-treatment synovial CD79a+ B cells [39,40], pre-treatment synovial molecular signatures [35] and the reduction of synovial plasma cells [42] have been identified as factors associated with response to RTX. However, the small number of patients analysed and the use of different timepoints makes it difficult to draw conclusions on the association of synovial B cell signatures with treatment response to RTX.

Eight research articles were identified that specifically evaluated if inflammatory markers could predict a response to TOC [43-50] and none for other IL-6 inhibitors. Five were observational studies [44-46,48,49] and 3 post-hoc analysis of randomised controlled trials (RCTs) [43,47,50]. None of the post-hoc RCT analyses found an association between baseline levels of inflammatory markers and response. In particular, a pooled analysis of 5 RCTs including 4,186 participants did not find any associations between baseline CRP levels and DAS28 at 24 weeks. In one prospective observational study, higher quartiles of levels of CRP at baseline were associated with lower discontinuation of TOC [44].

A large number of additional exploratory studies have searched for biomarkers linking changes in blood or synovium with response to RTX or TOC in RA. Abnormal B cell clonality has been shown to play an important role in RA pathogenesis [51,52] and other immune-mediated inflammatory diseases (IMID) [53]. While CD20+ B cells are virtually completely depleted in peripheral blood for up to 24 weeks following treatment with RTX, depletion of CD20+ B cells in synovium is patchy and incomplete in some individuals [cite]. The importance of persistent circulating plasmablasts / pre-plasma cell precursors which are largely CD20-ve in mediating treatment resistance to RTX has been reported using high-sensitivity FACS [24,28,54,55]. In comparison, blood memory B cell phenotype has also been linked to differential response to TOC [56]. Persistence of dominant B cell clones in both blood and synovium following treatment with RTX correlated with poor clinical response to RTX [57]. This effect has been presumed to be due to lack of depletion of CD20-ve plasmablast populations in blood and synovium. While ectopic lymphoid structures (ELS) in the synovium have been shown to support full development of pathogenic long-lived anti-CCP antibody secreting plasma cells [58], it is unknown whether post-RTX residual B cells or plasma cells resident in synovial tissue have sufficient local inflammatory drive to sustain inflammation and autoimmunity through early B cell repopulation and continuous production of autoantibodies. We hypothesise that the most fully functional synovial ELS may be capable of early regeneration of pathogenic B clones in situ and act as survival niches for long-lived autoreactive plasma cells.

In addition to B and plasma cells, other lineages have been also shown to be important in RA pathogenesis and response to therapy. For example, synovial macrophages have been shown to influence disease progression [59], remission [60] and treatment response [61]. RA specific fibroblast populations identified using single cell RNA-seq [62,63] have also been shown to drive RA immunopathogenesis [64]. We have shown that the molecular pathology of RA is a highly

heterogenous demonstrating in the early disease (MRC Pathobiology of Early Arthritis Cohort: PEAC), prior to DMARD exposure, the existence of 3 synovial pathotypes: Lympho-myeloid (B-cell rich ectopic lymphoid structures: ELS+ve), Diffuse-Myeloid (prevalence of sublining macrophage infiltration but B-cell poor), or pauci-immune/fibroid (prevalence of stromal cells but poor in B-cells) [65-67]. Notably, though these histo-pathotypes are not pure distinct entities, they are associated with related partially overlapping molecular profiles, as determined by synovial RNA-seq, that has enabled a better characterisation of disease strata linked to clinical heterogeneity [66]. For example, the molecular pathology of the synovial tissue has enabled the identification of specific signatures associated with more severe disease outcome (erosive load at 12 months) and treatment response to csDMARDs [65-67]. Of critical importance, at clinical follow-up, over 70% of PEAC patients with the pauci-immune/fibroid pathotype, displaying and FGF related signature, were non-responders to csDMARD [65]. Similarly, in an independent csDMARD-IR population, only 28.6% in the pauci-immune/fibroid pathotype responded to treatment with anti-TNF therapy [68]. Consistent with this, the preliminary RNA-Seq analysis in R4RA described above also identified a fibroblastic growth factor signature associated with impaired response to RTX (Fig. 2A).

Blood markers have also been identified which link to changes in synovial cellularity including serum CXCL13 which is linked to synovial B cell infiltration, soluble ICAM1 (sICAM1) which is associated with the presence of tissue macrophages [69], and blood interferon signature which is associated with the presence of synovial B cell and plasma cells [66]. Therefore, based on the published literature and our own observations, diverse cellular and molecular pathways could be involved in the mechanisms of response and non-response of RTX and TOC. In this application, capitalising on samples/datasets from the R4RA study, we aim to elucidate such mechanisms as described in detail below.

Hypotheses

Having identified the current unmet needs and the still unanswered research questions, capitalising on samples/ datasets from the R4RA study, we will define mechanisms of response and non-response to RTX and TOC by addressing the following hypotheses:

Treatment-specific processes leading to resistance to individual drugs are characterised by:

i) *Low or absent expression of the RTX and TOC targets* in RA synovium: e.g. a) pauci-immune/fibroid pathotype and/or b) low CD20 and IL-6/IL6-R pathways and/or c) critical downstream cells and signalling pathways (IL-6 receptor, cofactor gp130 and IL-6R/gp130 adaptors), leading to failure of response to RTX or TOC or both.

ii) *Incomplete depletion of synovial CD20+ B cells following RTX*, leading to early circulating activated CD27+ memory B cell subset repopulation as detected by hsFACS, which may correlate with reactivation of disease in between RTX infusions.

iii) *CD20 negative B cell compartment (plasma cells, plasmablasts) may mediate resistance to RTX and/or TOC.*

iv) *Persistence of blood and/or synovial CD27+ memory B cells* representing ongoing hyper-B cell differentiation drive, correlating with reduced response to RTX and/or TOC.

v) *Blood markers* (IFN signature, JCHAIN transcript, chemokines) representing active synovial processes, which may aid prediction of responsiveness to specific therapies.

These studies, by examining treatment-specific mechanisms of failure for each individual drug RTX & TOC, will inform further investigations attempting to characterise the mechanisms underlying the refractory RA phenotype.

4. Aims and Objectives

The original NIHR funding supported the delivery of the R4RA clinical trial, basic immunostaining and RNA-sequencing. In this application we will expand our investigation to test the above hypotheses through the delivery of the following objectives:

1. Elucidate whether specific synovial cell lineages and histo-pathotypes are linked to RTX or TOC response/non-response. Specific synovial B/T cell, fibroblast and macrophage subsets will be correlated with response/ non-response to RTX and/or TOC.
2. Determine whether diverse B cell differentiation states are responsible for the mechanisms of response to RTX or TOC [57]. We will perform deep immunophenotyping of synovial biopsies pre- and post-treatment and high sensitivity B cell subset flow cytometry of peripheral blood at multiple time points. Incomplete depletion in synovium, persistence of CD20-ve populations especially long-lived plasma cells, and repopulation of B cell subsets post-RTX, notably activated CD27+ memory B cells and/or plasmablasts will be correlated with response/ resistance to RTX or TOC therapy. We will determine whether pathogenic B cell subsets derive from synovial disease tissue or are newly spawned from bone marrow or SLOs.
3. Determine whether pathogenic B cell clonotypes are responsible for the mechanisms of non-response to RTX or TOC [57]. We will perform BCR sequencing of synovial biopsies. Persistence or repopulation of specific B cell clones will be correlated with response/ resistance to RTX or TOC therapy. Family tree analysis of B cell clone repopulation in blood versus synovium will be used to determine whether pathogenic B cell clones derive from synovial disease tissue niche or are newly spawned from bone marrow or SLOs.
4. Define whether RNA-Seq drug target gene modules and response prediction models derived from R4RA can be validated in STRAP.
    a. Low/high expression of IL-6 and IL-6 receptor gene module expression and/or the above B-cell signatures will be correlated with poor/good response to TOC in R4RA and compared with STRAP.
    b. FANTOM5 B cell gene module and RTX response prediction 35-gene model previously identified in R4RA RNA-Seq data will be validated in STRAP RNA-Seq data and meta-analysed.
5. Test the hypothesis whether previously reported biomarkers in the peripheral blood act as surrogates of tissue pathology (akin to the concept of a ‘liquid biopsy’) including serum CXCL13, sICAM1 [69] and blood type I interferon (IFN) response gene signature and/or treatment response including CCL19, IL-33, free light chains and IgJ. Thus, we will evaluate whether the above blood signatures are mechanistically linked to response to RTX and TOC, and whether their modulation heralds response to treatment and/or remission vs resistance.

Deliverables
Months 1–9 Milestone 1: R4RA RNA-Seq analysis and validation of prediction models in STRAP
Months 1–10 Milestone 2: BCR sequencing analysis
Months 2–5: Luminex serological analysis
Months 5–8: Blood qPCR analysis
Months 8–12 Milestone 3: Integrate serological/transcript predictors of pathotype & response
Months 3–16: Measure plasmablast/pre-plasma cell repopulation by hsFACS
Months 3–22: Multiplex IHC followed by Hyperion in selected samples for deep immune-histological phenotyping of synovial tissue
Months 12–15 Milestone 4: Integrate IL-6 drug target module analysis with B cell clonotype and B cell subset analysis.
Months 15–22 Milestone 5: Full clonotype/phenotype integration.
Months 22-24: Full report submitted to NIHR.

5. Research Plan / Methods

Research design
The proposed research will utilise the advanced cellular and molecular techniques to evaluate the mechanisms of action of RTX and TOC in matched blood and synovial biopsies pre- and post-treatment which arose from the R4RA trial. The RCT design of the original trial will ensure that robust clinical data is available to contextualise the biological results. In order to identify the drug-refractory endotype, use of cross-over RCT data from R4RA will be employed. In R4RA, 37

patients failed both biologics representing a group of truly drug-recalcitrant patients (who by definition also failed standard DMARDs and anti-TNF prior to study entry). This group will be compared against patients, including cross-over patients, who responded to each biologic.

Importantly, from the above literature review it is clear that abnormal B cell clonality has been shown to play an important role in RA pathogenesis [51,52] and other immune-mediated inflammatory diseases (IMID) [53]. Persistent low levels of circulating plasmablasts / pre-plasma cell precursors detected by high-sensitivity FACS have been associated with resistance to RTX [24,28,54,55]. Although RTX profoundly depletes CD20+ B cells for up to 24 weeks, persistence of specific clones thought to derive from CD20-ve plasmablasts correlates with poor clinical response to RTX [57]. Prior studies have shown that synovial macrophage filtration influences disease progression [59] and treatment response [61]. RA specific fibroblast populations identified using single cell RNA-seq [62,63] have been shown to drive RA immunopathogenesis [64]. Our own studies described above have also demonstrated that the molecular pathology of the synovial tissue is associated with more severe disease outcome (erosive load at 12 months) and treatment response to csDMARDs [65-67]. Of critical importance, at clinical follow-up, over 70% of PEAC patients with the pauci-immune/fibroid pathotype, displaying and FGF related signature, were non-responders to csDMARD [65]. Similarly, in an independent csDMARD-IR population, only 28.6% in the pauci-immune/fibroid pathotype responded to treatment with anti-TNF therapy [68]. Therefore, in this proposal we will investigate whether synovial pathotypes are also responsible for response or lack of response to RTX, TOC or both (refractory RA).

Multiple blood markers, including serum markers CCL19, IL-33 and free light chains [33,34,69,70], and blood RNA markers including interferon gene signature [71,72] and IgJ [29], have previously been correlated with therapeutic response to RTX in RA. In addition, CXCL13 and soluble ICAM1 (sICAM1) in the circulation have been reported to be surrogate biomarkers of the lymphoid (B-cell rich) and myeloid (B-cell low) prevalent pathotypes respectively [69], raising the tantalising prospect of defining tissue pathology by a blood “liquid biopsy”. Notably in the same study, when participants in the ADACTA trial (comparing effectiveness of the anti-TNF agent Adalimumab v TOC) were stratified according to low CXCL13 and high sICAM1 (myeloid pathotype) response rates favoured Adalimumab (ACR50: 42% v 20% respectively), whereas in patients with high CXCL13 and low sICAM1 (lymphoid pathotype) response rates dramatically favoured TOC (ACR50: 13% v 69% respectively) [69]. These results are consistent with the notion that, in addition to being a powerful pro-inflammatory cytokine, IL-6 is also a potent factor for B cell growth and differentiation [9-11], providing a plausible explanation for the powerful effect in patients with a presumed lymphoid pathotype. Therefore, in this proposal we will also investigate whether, peripheral blood biomarkers can act as surrogate of disease pathology (liquid biopsy) and/or treatment response.

Aim 1: *To elucidate whether specific synovial cell lineages and histo-pathotypes are linked to RTX or TOC response/non-response*
QuPATH histology digital image analysis of cell lineages in synovium. Whole slides of synovial biopsies from R4RA patients at baseline and 16 weeks have previously been stained by Haematoxylin & Eosin and Immunohistochemistry (IHC) (CD20+ B-cells, CD79a+ B-cells, CD3+ T cells, CD68+ Macrophages, CD138+ Plasma-cells) as part of the R4RA study. In this proposal, slides will be stained for additional B-cell and plasmablast markers as well as ELS markers, as described in details in Aim 2. Furthermore, here we will analyse by multiplex IHC T cells (CD4, CD8, ICOS, PD1, HLA-DR), fibroblast (Thy-1, CD34, HLA-DR) and macrophage (CD68, CD163, CD206, TREM2, MERTK, SPP1, S100A12, HLA-DR) among other markers. Slides will be analysed with open source software QuPath [73], using an in-house developed script for automated Digital Image Analysis (DIA), which has already been applied to CD20 IHC staining [74]. Following tissue recognition, the positive cell detection function will be used to set thresholds of staining intensity and recognise positive cells. As opposed to the semi-quantitative scores used so far, this will provide a quantitative measurement (cell density), similar to DIA-derived scores developed in oncology [75], which will be correlated against treatment response to RTX or TOC, comparing each drug against the other. Quantitative histology results for each cell type will also be correlated with cell-specific transcript levels in the synovial RNA-Seq and with treatment response.

Finally, we will investigate the relationship of the 3 histo-pathotypes described above and treatment response.
Expected outcome. This investigation will determine whether specific subsets of B/T cells fibroblasts or macrophages, or presence of plasma cells in synovium at baseline are differentially associated with treatment failure for RTX and/or TOC.

Aim 2: *To determine whether diverse B cell differentiation states are responsible for the mechanisms of response to RTX or TOC.*
i) Deep phenotyping of B cell subsets including CD20 negative plasmablasts/plasma cells and pro-B cell maturation immune cell types in synovium. In parallel and linked to the BCR analysis (aim 3, below), deep histological phenotyping of the corresponding synovium will be performed using Hyperion imaging mass cytometry in selected samples on the basis of the results of Multiplex IHC. This will allow up to 37 markers per cell to be labelled in order to identify which subsets are affected and/or resistant to specific drug treatment with RTX or TOC, using heavy metal labelled antibody markers for B cell memory subsets (CD20, CD79, IgM, IgD, CD38, CD27, HLA-DR), plasma cell (CD138), T cell (CD4, CD8, ICOS, PD-1, HLA-DR, Granzyme B & K), macrophage (CD68, CD163, CD206, TREM2, MERTK, SPP1, S100A12, HLA-DR), follicular dendritic cell (CD21, CNA42), and fibroblast subpopulations (Thy-1, CD34, HLA-DR). Prior studies have shown that synovial macrophage infiltration influences disease progression [59] and treatment response [61]. RA specific fibroblast populations identified using single cell RNA-seq [62,63] have been shown to drive RA pathogenesis [64]. Similarly to aim 1, Hyperion results for each cell type will be compared with cell-specific transcript levels in synovial RNA-Seq and correlated with treatment response.
Expected outcome. This investigation will elucidate whether specific subsets of memory B cells and/or other pathogenic cell populations e.g. infiltrating tissue macrophages, Thy-1+ fibroblasts drive persistence of inflammation in the synovium leading to treatment failure for RTX or TOC. Comparison of biopsies at baseline and 16 weeks will show whether persistence of specific cell types, specifically incomplete depletion of synovial B cells post-RTX, or persistence of specific cell types at 16 weeks is associated with treatment failure for RTX and/or TOC and whether there are contrasting drug-specific cellular differences which correlate with response/non-response.
ii) Measuring repopulation of plasmablasts/ pre-plasma cell precursors by high sensitivity flow cytometry. B cell depletion and repopulation in periphery will be examined by FACS by 10-15 colour flow cytometry using markers CD19, CD20, CD24, CD27, CD38, IgD, IgA, IgM and exclusion markers CD3, CD14, CD56, live/dead stain. The importance of persistent circulating plasmablasts / pre-plasma cell precursors in mediating treatment resistance to RTX has been reported using high-sensitivity FACS [24,28,54,55].
Expected outcome. This investigation will identify the time to first, low-level repopulation of circulating plasmablasts and whether early repopulation mediates resistance to RTX and whether persistence of circulating B cell subtypes mediates resistance to TOC. Importantly, peripheral blood results will be compared and contrasted with the synovial histopathological analyses described above.

Aim 3: *Mapping B cell clonality in tissue and peripheral blood as a mechanism of non-response to RTX and TOC.* The clonal relationship between B cells infiltrating the synovium at baseline and those found in the post-treatment biopsy will be assessed using B cell receptor repertoire and family tree analysis as previously described [51]. Diversity of the peripheral blood and synovial tissue BCR repertoire pre and post-treatment will be examined as a previous study has shown that the degree of depletion of the diversity of the B cell repertoire is correlated with depth of response to RTX [57]. RNA will be processed with a 5'RACE reaction using 3'primers that bind to the constant regions of the IgM, IgD, IgG, IgA, IgE, Igκ and Igλ locus. The designed 3'-primers binds to the antibody constant CH1/CL regions, which allow to amplify the variable regions (VH/VL), and a portion of the constant regions. The resulting fragments will be subject to NGS library preparation and sequencing on the Illumina MiSeq (2x300bp), 2M Paired End reads/sample. B cell repertoire analysis will include CDR 1-3 nucleotides sequences, amino acid sequences, relative abundances; VDJ distribution; CDR3 length distribution, frequency analysis and clustering and isotype usage. To account for greater number of BCR RNA in plasmablasts and plasma cells, the percentage of reads /isotype and the normalized isotype use will reduce potential biases from

differential RNA per cell. The network generation algorithm and network properties will be calculated using previously described methodology [76]. Pre/post treatment family tree analysis comparing synovium and blood will be employed to determine whether clonotypes are repopulated locally from residual B cell clones in synovium or newly derived clones for example from the bone marrow and/or SLOs, and compared with therapeutic response and anti-CCP antibody (ACPA) status.
Expected outcome. This investigation will elucidate whether residual pathogenic CD20-ve plasma cell clones which are left untouched by RTX and/or TOC (not because of specific clonality but for example in relation to PK/PD pharmacology reasons) lead to persistently active disease; or whether new clones spawned from bone marrow or SLOs leads to treatment failure.

Aim 4: *Correlate drug targets and linked signalling apparatus with response to TOC and RTX.*
i) To investigate the mechanisms of action of TOC, we will quantify IL-6 drug target gene module using RNA-seq data for synovial expression of *IL6*, *IL6R* and its co-receptor *IL6ST* (gp130) and proteins directly linked to IL-6 receptor and gp130 signalling (including transcripts for STAT/JAK proteins). Using a bioinformatic analysis approach similar to weighted gene correlation network analysis (WGCNA) [77,78], integrating information from the STRING protein-protein interaction database (https://string-db.org), the IL-6 drug target module will be correlated with differential response to TOC vs RTX. CD20 and downstream signalling pathways and transcriptional regulation will be investigated using the same methodology.
ii) FANTOM5 B cell gene RNA-Seq module and RTX response 35-gene prediction model will be validated in the STRAP study using RNA-Sequencing currently being performed on synovial biopsies (funded by MRC, data available Sept 2020). Statistical analysis of the optimal cut-off for each model will be compared between R4RA and STRAP. Further analysis using response as a continuous outcome to improve power will also be performed. Meta-analysis pooling both studies will be performed to refine both models and compare prediction accuracy.
Expected outcome. This investigation will determine whether low or absent expression of IL-6, IL-6 receptor or its receptor-JAK-STAT signalling apparatus mediates treatment failure to TOC. Informatic analysis will determine whether the FANTOM5 B cell module or R4RA RTX prediction model are validated in STRAP and which model is most accurate for predicting response.

Aim 5: *Define surrogate blood markers linked to synovial pathology to elucidate mechanisms of drug response*
i) *Luminex analysis*. Serum markers CXCL13, CCL19, soluble ICAM1 (sICAM1), IL-33 and free light chains, which have previously been correlated with therapeutic response in RA [33,34,69,70], and up/downstream factors associated with B cell maturation/plasma cell differentiation and IL-6 secretion including IL-17, IL-21, lymphotoxin-β [79] will be assayed by Luminex pre- and post-treatment and correlated with clinical response and anti-CCP antibody levels, as well as incorporating data from B cell repopulation in peripheral blood in RTX treated individuals from Aim 2.
Expected outcome. This investigation will compare and contrast different serum markers associated with pathological changes in synovial tissue to determine which cytokines/chemokines are most important for driving B cell differentiation and tissue infiltration in RA, and thus whether high levels of these cytokines/chemokines are responsible for the mechanism of action or treatment failure to RTX or TOC.
ii) *qPCR analysis.* Type I IFN response signature in blood is associated with synovial B cell infiltration, ACPA production and predicts bone erosion in early RA [66]. IFN response signature [71,72] and B cell maturation associated blood transcripts will be measured using Taqman realtime PCR employing a limited geneset and housekeeping genes, in addition to measuring IgJ mRNA [29]. qPCR on whole blood will be performed pre and post treatment and correlated with response.
Expected outcome. This investigation will elucidate the role of type I IFN response in synovial B cell filtration and whether persistent aberrant blood IFN response drives B cell maturation leading to failure of RTX or TOC.

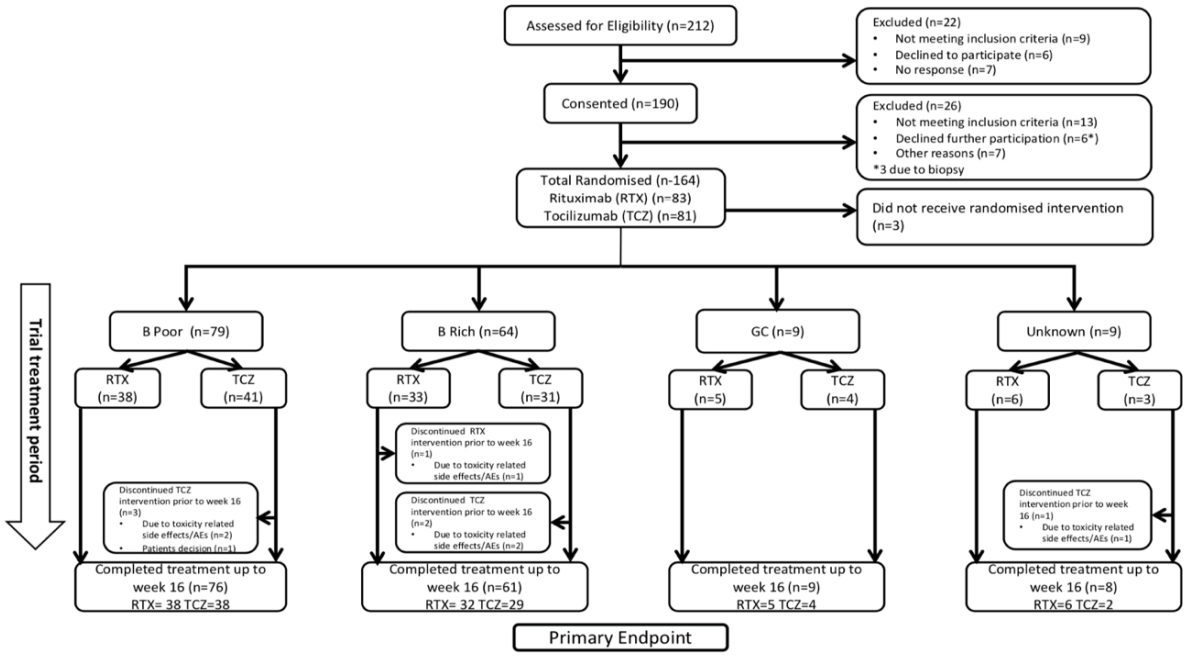
Study Population

Design: We conducted a phase IV open label, multi-centre, randomised control trial (RCT). Patients were randomised to receive RTX or TOC stratified according to histological classification of baseline synovial biopsy (B-cell poor, B-cell rich, GC+ or unknown) and by site (QMUL vs all other sites). Patients were followed up at 4 weekly intervals throughout the 48 week trial treatment period where RA disease activity measurements and safety data were collected. An optional repeat synovial biopsy of the same joint sampled at baseline was performed at 16 and 48 weeks.

Setting: Rheumatology outpatient clinics in 19 European centres.

Participants: Patients aged 18 years or over, fulfilling 2010 ACR/EULAR classification criteria for RA who were eligible for treatment with RTX therapy according to UK NICE guidelines (failing or intolerant to csDMARD therapy and at least one anti-TNF agent (excluding trial IMPs) were eligible for recruitment to the study and identified through rheumatology outpatient clinics at each study site. All patients eligible to take part in research were offered the same opportunity to take part in the research regardless of age, gender, sexual orientation, marital status, disability, ethnicity, socioeconomic status, access to healthcare.

Patients and demographics: 212 patients were screened, 190 were consented and 164 patients underwent randomisation (see CONSORT patient flow diagram below). The trial ended as recruitment targets were reached. 83 patients were randomised to receive treatment with RTX and 81 with TOC. 161 patients received investigational medicinal product (IMP). Baseline characteristics, disease activity and histological groups were balanced across the treatment arms. Most patients were female (80%) and the majority were sero-positive for rheumatoid factor (67%) or anti-citrullinated peptide antibodies (77%). Median disease duration was 9 years (IQR 4,19). Disease activity was high [mean DAS28-ESR 5.8 (SD 1.2)]. 49% patients were classified as B-cell poor and 40% as B-cell rich.



Proposed deliverables

1. Months 1 – 9 (01/04/21 – 31/01/22) Milestone 1: In silico analysis of synovial RNA-Seq data to determine if expression of IL6/IL6R downstream genes predict response to TOC; validation of R4RA prediction models in STRAP, refinement of model cut-offs and meta-analysis; draft paper.
2. Months 1 – 10 (01/04/21 – 30/02/22) Milestone 2: BCR profiling of synovial and blood RNA (sequencing outsourced) followed by clonotype analysis and family tree analysis.

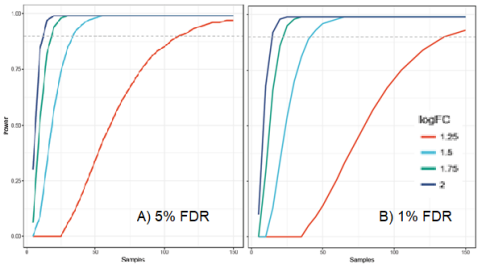
3. Months 2 – 5 (01/05/21 – 28/10/21) Serological analysis of predicted blood biomarkers of synovial pathotype and response to RTX or TOC by Luminex.
4. Months 5 – 8 (01/10/21 – 31/01/22) Transcriptomic analysis of predicted blood biomarkers of synovial pathotype and response to RTX or TOC by qPCR or nanoString autoimmune panel.
5. Months 8 – 12 (01/01/22 – 31/04/22) Milestone 3: Integration of serological and transcriptomic biomarkers as predictors of pathotype and/or response to RTX or TOC; draft paper.
6. Months 3 – 16 (01/08/21 – 31/08/22) Measuring repopulation of plasmablasts/ pre-plasma cell precursors by high sensitivity flow cytometry on frozen PBMC samples (baseline, 3 time-point post-treatment).
7. Months 3 – 22 (01/03/22 – 30/02/23) Deep histological phenotyping of synovial tissue by multiplex IHC, and on selected samples Hyperion imaging mass cytometry.
8. Months 12 – 15 (01/05/22-31/07/22) Milestone 4: Integrate IL-6 drug target module analysis with B cell clonotype and B cell subset analysis to elucidate mechanisms of failure of TOC.
9. Months 15 – 22 (01/09/22 – 31/04/23) Milestone 5: Full clonotype and phenotype (PBMC & synovium) integration; draft paper.

Sample size

For the original R4RA trial, a sample size of 82 B-cell-poor patients was planned to provide 90% power to detect a 35% (assuming 55% response in TOC and 20% response in RTX) difference in the proportion of patients who had a response. The assumed proportions of B-cell-poor, B-cell-rich and GC+ recruited patients were 60%, 35% and 5% respectively. After accounting for 10% ungradable biopsy samples and a 5% dropout rate, it was estimated that a total of 160 patients was required to achieve 90% power for the study. No power calculation was conducted for the B-cell-rich population.

The mechanistic laboratory studies proposed in this application are by their nature exploratory and pilot data are generally either not available or small scale, thus rendering sample size and power calculations impossible. However, previous studies examining clonotype diversity following RTX demonstrated significant results (P<0.01) in a comparison of 5 non-responders and 9 moderate responders following treatment with RTX [57]. Thus, we anticipate that planned analyses of paired synovial biopsies at baseline pre-treatment and week 16 will have sufficient samples to test hypotheses in the proposed analyses. From R4RA we have synovial biopsies from all patients at baseline (n=161), all of which have already been RNA-Sequenced. Paired pre/post treatment synovial and blood samples are available on the following: RTX non-responders n=26, responders n=15; TOC non-responders n=14, responders n=10.

Power and Sample Size Exemplar: Transcriptomics



Using the method of Guo et al [80] and parameters derived from PEAC synovial RNAseq [66], we estimate sample size to achieve 90% power based on the observation that 1-5% of transcripts are likely to be prognostic of response, depending on response measure. We observed an average read count of ~100 in prognostic genes prior to normalisation; and a minimum log fold change of approximately 1.4 after modelling; and a global dispersion estimate of 0.137, as estimated by empirical Bayes procedure [81]. Imposing a 5% FDR threshold and a target log fold change of 1.5, we find that a study would require 35 subjects (fig. A) to achieve 90% power to identify transcriptomic predictors of response. At a more stringent 1%, we can maintain 90% power with 40 subjects (fig. B). Power curves projected across an expected range of fold changes at 1% and 5% differential expression are shown in the enclosed figure. Predicted power is anticipated to be similar for the range of analytical contrasts proposed.

For learning to predict from high-dimensional biomarker panels, an alternative approach to calculating required sample size is described by McKeigue [82]. This is based on calculating the expected information for discrimination of the predictive model learned from a training sample, as a fraction of the information that would be obtained with the optimal model that would be learned from a sample of infinite size. We calculate that if about 1% of 10,000 gene transcripts have nonzero effects, and if the optimal predictive model has a C-statistic of 0.9, a sample size of about

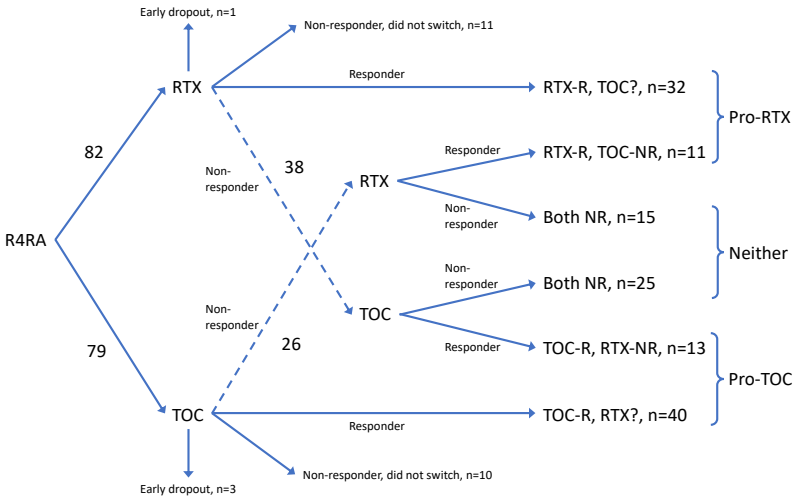
200 cases (cases/variable ratio 0.02) is enough to learn a model that extracts about half the information for discrimination that an optimal model would have. In practice we can exploit prior information to reduce the required sample size: for instance, preselecting genes by pathway or by their previously identified associations with endotypes.

Please see "**9. Ethics / Regulatory Approvals**". The patient consent form for the R4RA clinical trial specifies that samples will be transferred to the HTA and ethically approved EMR Biobank REC #17-WS-0172 for further research which is reviewed by the EMR Biobank oversight management committee. This allows research to continue beyond 31 May 2019.

Statistical Analysis

The statistical analysis can be summarised as two main categories of analysis: i) hypothesis testing for statistical differences between observed experimental populations; ii) predicting response to treatment and selection of biomarkers from high-dimensional data. For i), standard statistical approaches using will be applied. Notably the outcome (drug response) will be modelled with multiple categorical endpoint measures (CDAI 50% improvement, CDAI-MTR, EULAR DAS28-ESR/CRP response, Sharp van der Heijde score X-ray progression etc) as described in the original R4RA statistical analysis plan (SAP). Statistical testing of experimental variables will be performed by calculating the relative risk ratio (RR) comparing both response vs non-response for each drug as well as comparing the efficacy (response vs non-response) for both drugs simultaneously in the cohort stratified by test variable, e.g. B cell poor individuals, in a similar manner to the original statistical methodology for the R4RA trial. For contingency table tests, uncorrected chi square test ($\geq$5 individuals in all classes) or Fisher's exact test (<5 individuals in any class) will be used. Accuracy of binary outcome response predictors will be compared using area under receiver operating characteristic (ROC) curve analysis for each drug. When necessary, logistic regression models will be employed to enable incorporation of covariates and to analyse for interaction between predictors. Secondary analyses will be performed using continuous response endpoints specified in the R4RA SAP (e.g. CDAI, DAS28-ESR, DAS28-CRP, total ultrasound joint scores, total Sharp X-ray score etc), with addition of the more recent DAS-2C [83]. Experimental variables will be tested against continuous response outcomes for each drug separately using linear regression, as well as testing for drug interaction by likelihood ratio test to contrast difference in effect between RTX and TOC. Sharp X-ray score, which as a form of count data is typically log distributed, will be tested by Poisson regression. Longitudinal data, e.g. hsFACS B cell repopulation analyses, will be tested using mixed effects linear regression models using the lme4 package. B cell receptor repertoire and family tree analyses will be performed as previously described [51], applying network analysis methods [76].

As an expansion to the analyses described above comparing responders vs non-responders in each drug cohort, cross-over RCT data from R4RA will be employed to identify a multidrug-refractory endotype: in R4RA, 37 patients failed both biologics representing a group of truly drug-recalcitrant patients who have now failed both RTX & TOC, and prior to study entry failed standard DMARDs and anti-TNF. This refractory group will be compared against patients, who responded to each biologic in 3-way analyses (flow chart, below).



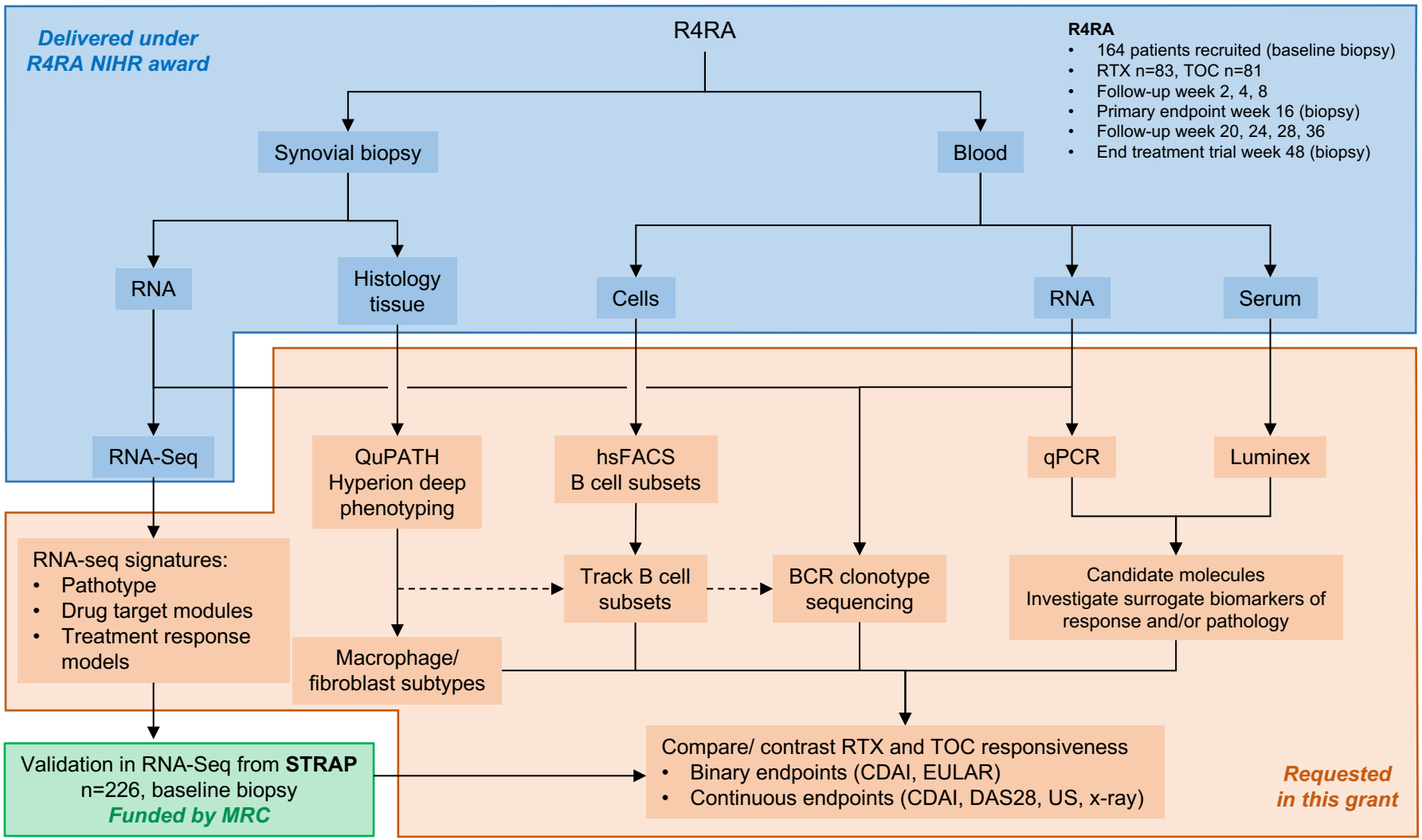
For high dimensional analyses in ii), leading-edge methods for testing prediction of outcome and selection of biomarkers from multi-dimensional panels will also be employed, by first building predictive models (Bayesian regression with hierarchical shrinkage priors to control overfitting) and evaluating predictive performance by cross-validation, then using projection predictive variable selection to select the most predictive biomarkers (ongoing collaboration with Prof Paul McKeigue, Edinburgh, through the MATURA consortium project). For analysis of RNA-Seq data, read count data will be fitted using negative binomial regression (the standard approach for overdispersed count data) using a custom R package written by the Lewis group built around the lme4 and MASS R packages allowing analysis of read count matrices using negative binomial regression and fully flexible model formulae including interaction and/or mixed effects models for longitudinal read count data. Multi-gene predictive models will be developed using either normally distributed transformed data (DESeq2 variance stabilising transform) within a generalised linear model, and compared with library size adjusted TMM read counts modelled with negative binomial regression predictive models. Outcome will be modelled as either categorical response by logistic regression, or as continuous response as a generalised linear model using the clinical outcome measures listed above. Different machine learning techniques will be compared for variable selection including LASSO/elastic net regression (using glmnet) and random forest. Predictive performance will be evaluated on test data by cross-validation using the caret package in R, with the same cross-validation folds used each time to maintain consistency in model comparisons. Models derived by glmnet and random forest will be compared with models derived through the rstanarm package which uses Bayesian estimation and MCMC to derive a sparse model with biomarker coefficients learned from the data. For binary outcomes, model accuracy and C-statistic will be reported. Where appropriate, model comparison will be based on differences in test log-likelihood [84]. Predictive models trained on the R4RA cohort will be validated on synovial RNA-Seq data from the independent STRAP cohort.

6. Ethics / Regulatory Approvals

The project will be carried out in compliance with all relevant regulations: Research Ethics; the Human Tissue Act and the General Data Protection Regulation (GDPR). Samples and data used in this project are from the R4RA Clinical Trial. The trial was approved in 2012 by the MHRA and the Research Ethics Committee for Wales, reference 12/WA/0307, with an end date of 31 May 2019. The patient consent form for the R4RA clinical trial specifies that samples will be transferred to the HTA and ethically approved EMR Biobank REC #17-WS-0172 for further research which is reviewed by the EMR Biobank oversight management committee. This allows research to continue beyond 31 May 2019.

No benefits or risks to patients, from whom the samples for this research have already been taken, have been identified.

Flow Chart: Plan of Investigation



Hypothesis:
Predictors of poor response to RTX
- Low synovial B cell infiltrate
- Incomplete depletion of synovial B cells
- Early blood CD27⁺ memory B cells repopulation
- Specific clonotypes in synovium driving early B cell repopulation

Hypothesis:
Predictors of poor response to TOC
- Low synovial IL6 signaling system expression
- Presence/persistence of synovial/blood CD27⁺ memory B cells
- Persistent pathogenic clonotypes in synovium

Hypothesis:
Predictors of poor response to both RTX/ TOC
- Synovial fibroblast/ macrophage subtypes
- Synovial fibroid pathotype
- Persistent synovial long-lived plasma cells

References

1. Humby F, *et al.* Rituximab versus tocilizumab in anti-TNF inadequate responder patients with rheumatoid arthritis (R4RA): 16-week outcomes of a stratified, biopsy-driven, multicentre, open-label, phase 4 randomised controlled trial. *Lancet* 2021; 397(10271):305-17. PMID: 33485455.
2. Safiri S, *et al.* Global, regional and national burden of rheumatoid arthritis 1990-2017: a systematic analysis of the Global Burden of Disease study 2017. *Ann Rheum Dis* 2019; 78(11):1463-71. PMID: 31511227.
3. Smolen JS, *et al.* Rheumatoid arthritis. *Lancet* 2016; 388(10055):2023-38. PMID: 27156434.
4. Kristensen LE, *et al.* Predictors of response to anti-TNF therapy according to ACR and EULAR criteria in patients with established RA: results from the South Swedish Arthritis Treatment Group Register. *Rheumatology (Oxford)* 2008; 47(4):495-9. PMID: 18316338.
5. Cuppen BV, *et al.* Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability. *Rheumatology (Oxford)* 2016; 55(5):826-39. PMID: 26715775.
6. Pitzalis C, *et al.* New learnings on the pathophysiology of RA from synovial biopsies. *Curr Opin Rheumatol* 2013; 25(3):334-44. PMID: 23492740.
7. Humby F, *et al.* Use of Ultrasound-Guided Small Joint Biopsy to Evaluate the Histopathologic Response to Rheumatoid Arthritis Therapy: Recommendations for Application to Clinical Trials. *Arthritis Rheumatol* 2015; 67(10):2601-10. PMID: 26097225.
8. Kelly S, *et al.* Ultrasound-guided synovial biopsy: a safe, well-tolerated and reliable technique for obtaining high-quality synovial tissue from both large and small joints in early arthritis patients. *Ann Rheum Dis* 2015; 74(3):611-7. PMID: 24336336.
9. Eto D, *et al.* IL-21 and IL-6 are critical for different aspects of B cell immunity and redundantly induce optimal follicular helper CD4 T cell (Tfh) differentiation. *PloS one* 2011; 6(3):e17739. PMID: 21423809.
10. Tanaka T, *et al.* IL-6 in inflammation, immunity, and disease. *Cold Spring Harbor perspectives in biology* 2014; 6(10):a016295. PMID: 25190079.
11. Jourdan M, *et al.* IL-6 supports the generation of human long-lived plasma cells in combination with either APRIL or stromal cell-soluble factors. *Leukemia* 2014; 28(8):1647. PMID: 24504026.
12. Giacomelli R, *et al.* Guidelines for biomarkers in autoimmune rheumatic diseases - evidence based analysis. *Autoimmunity reviews* 2019; 18(1):93-106. PMID: 30408582.
13. Oxford Centre for Evidence-based Medicine – Levels of Evidence (March 2009). Oxford: Centre for Evidence-Based Medicine, 2009. https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/ (*Accessed 26 Feb 2020*).
14. Breedveld F, *et al.* Rituximab pharmacokinetics in patients with rheumatoid arthritis: B-cell levels do not correlate with clinical response. *J Clin Pharmacol* 2007; 47(9):1119-28. PMID: 17766699.
15. Cohen SB, *et al.* Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: Results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial evaluating primary efficacy and safety at twenty-four weeks. *Arthritis Rheum* 2006; 54(9):2793-806. PMID: 16947627.
16. Cohen SB, *et al.* Continued inhibition of structural damage over 2 years in patients with rheumatoid arthritis treated with rituximab in combination with methotrexate. *Ann Rheum Dis* 2010; 69(6):1158-61. PMID: 20439295.
17. Edwards JC, *et al.* Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. *N Engl J Med* 2004; 350(25):2572-81. PMID: 15201414.
18. Emery P, *et al.* Efficacy and safety of different doses and retreatment of rituximab: a randomised, placebo-controlled trial in patients who are biological naive with active rheumatoid arthritis and an inadequate response to methotrexate (Study Evaluating Rituximab's Efficacy in MTX iNadequate rEsponders (SERENE)). *Ann Rheum Dis* 2010; 69(9):1629-35. PMID: 20488885.
19. Emery P, *et al.* The efficacy and safety of rituximab in patients with active rheumatoid arthritis despite methotrexate treatment: results of a phase IIB randomized, double-blind, placebo-controlled, dose-ranging trial. *Arthritis Rheum* 2006; 54(5):1390-400. PMID: 16649186.
20. Mease PJ, *et al.* Efficacy and safety of retreatment in patients with rheumatoid arthritis with previous inadequate response to tumor necrosis factor inhibitors: results from the SUNRISE trial. *J Rheumatol* 2010; 37(5):917-27. PMID: 20194448.
21. Rubbert-Roth A, *et al.* Efficacy and safety of various repeat treatment dosing regimens of rituximab in patients with active rheumatoid arthritis: results of a Phase III randomized study (MIRROR). *Rheumatology (Oxford)* 2010; 49(9):1683-93. PMID: 20463186.
22. Tak PP, *et al.* Inhibition of joint damage and improved clinical outcomes with rituximab plus methotrexate in early active rheumatoid arthritis: the IMAGE trial. *Ann Rheum Dis* 2011; 70(1):39-46. PMID: 20937671.
23. Brezinschek HP, *et al.* B lymphocyte-typing for prediction of clinical response to rituximab. *Arthritis research & therapy* 2012; 14(4):R161. PMID: 22770118.
24. Dass S, *et al.* Highly sensitive B cell analysis predicts response to rituximab therapy in rheumatoid arthritis. *Arthritis Rheum* 2008; 58(10):2993-9. PMID: 18821683.
25. Vancsa A, *et al.* Longterm effects of rituximab on B cell counts and autoantibody production in rheumatoid arthritis: use of high-sensitivity flow cytometry for more sensitive assessment of B cell depletion. *J Rheumatol* 2013; 40(5):565-71. PMID: 23547216.
26. Sellam J, *et al.* Blood memory B cells are disturbed and predict the response to rituximab in patients with rheumatoid arthritis. *Arthritis Rheum* 2011; 63(12):3692-701. PMID: 22127692.
27. Roll P, *et al.* Anti-CD20 therapy in patients with rheumatoid arthritis: predictors of response and B cell subset regeneration after repeated treatment. *Arthritis Rheum* 2008; 58(6):1566-75. PMID: 18512772.
28. Vital EM, *et al.* Management of nonresponse to rituximab in rheumatoid arthritis: predictors and outcome of re-treatment. *Arthritis Rheum* 2010; 62(5):1273-9. PMID: 20131284.
29. Owczarczyk K, *et al.* A plasmablast biomarker for nonresponse to antibody therapy to CD20 in rheumatoid arthritis. *Science translational medicine* 2011; 3(101):101ra92. PMID: 21937757.

30. Ruyssen-Witrand A*, et al.* Association between -871C>T promoter polymorphism in the B-cell activating factor gene and the response to rituximab in rheumatoid arthritis patients. *Rheumatology (Oxford)* 2013; 52(4):636-41. PMID: 23264555.
31. Quartuccio L*, et al.* The 158VV Fcgamma receptor 3A genotype is associated with response to rituximab in rheumatoid arthritis: results of an Italian multicentre study. *Ann Rheum Dis* 2014; 73(4):716-21. PMID: 23505228.
32. Fabris M*, et al.* The CC homozygosis of the -174G>C IL-6 polymorphism predicts a lower efficacy of rituximab therapy in rheumatoid arthritis. *Autoimmunity reviews* 2012; 11(5):315-20. PMID: 20974296.
33. Sellam J*, et al.* Serum IL-33, a new marker predicting response to rituximab in rheumatoid arthritis. *Arthritis research & therapy* 2016; 18(1):294. PMID: 27964756.
34. Sellam J*, et al.* CCL19, a B cell chemokine, is related to the decrease of blood memory B cells and predicts the clinical response to rituximab in patients with rheumatoid arthritis. *Arthritis Rheum* 2013; 65(9):2253-61. PMID: 23740460.
35. Hogan VE*, et al.* Pretreatment synovial transcriptional profile is associated with early and late clinical response in rheumatoid arthritis patients treated with rituximab. *Ann Rheum Dis* 2012; 71(11):1888-94. PMID: 22736099.
36. Kavanaugh A*, et al.* Assessment of rituximab's immunomodulatory synovial effects (ARISE trial). 1: clinical and synovial biomarker results. *Ann Rheum Dis* 2008; 67(3):402-8. PMID: 17644541.
37. Kormelink TG*, et al.* Decrease in immunoglobulin free light chains in patients with rheumatoid arthritis upon rituximab (anti-CD20) treatment correlates with decrease in disease activity. *Ann Rheum Dis* 2010; 69(12):2137-44. PMID: 20679475.
38. Rosengren S*, et al.* CXCL13: a novel biomarker of B-cell return following rituximab treatment and synovitis in patients with rheumatoid arthritis. *Rheumatology (Oxford, England)* 2011; 50(3):603-10. PMID: 21098574.
39. Teng YKO*, et al.* Immunohistochemical analysis as a means to predict responsiveness to rituximab treatment. *Arthritis Rheum* 2007; 56(12):3909-18. PMID: 18050222.
40. Teng YKO*, et al.* Residual inflammation after rituximab treatment is associated with sustained synovial plasma cell infiltration and enhanced B cell repopulation. *Ann Rheum Dis* 2009; 68(6):1011-6. PMID: 18647852.
41. Teng YKO*, et al.* Induction of long-term B-cell depletion in refractory rheumatoid arthritis patients preferentially affects autoreactive more than protective humoral immunity. *Arthritis research & therapy* 2012; 14(2):R57-R. PMID: 22409963.
42. Thurlings RM*, et al.* Synovial tissue response to rituximab: mechanism of action and identification of biomarkers of response. *Ann Rheum Dis* 2008; 67(7):917-25. PMID: 17965121.
43. Bay-Jensen AC*, et al.* Early changes in blood-based joint tissue destruction biomarkers are predictive of response to tocilizumab in the LITHE study. *Arthritis research & therapy* 2016; 18:13. PMID: 26787505.
44. Forsblad-d'Elia H*, et al.* Drug adherence, response and predictors thereof for tocilizumab in patients with rheumatoid arthritis: results from the Swedish biologics register. *Rheumatology (Oxford)* 2015; 54(7):1186-93. PMID: 25505001.
45. Hoshi D*, et al.* Disability is the major negative predictor for achievement of Boolean-based remission in patients with rheumatoid arthritis treated with tocilizumab. *Modern rheumatology / the Japan Rheumatism Association* 2013; 23(6):1205-10. PMID: 23456321.
46. Izumi K*, et al.* Baseline Serum Osteopontin Levels Predict the Clinical Effectiveness of Tocilizumab but Not Infliximab in Biologic-Naive Patients with Rheumatoid Arthritis: A Single-Center Prospective Study at 1 Year (the Keio First-Bio Cohort Study). *PloS one* 2015; 10(12):e0145468. PMID: 26698858.
47. Kato M*, et al.* Predictive value of serum amyloid a levels for requirement of concomitant methotrexate in tocilizumab initiation: A post hoc analysis of the SURPRISE study. *Modern rheumatology / the Japan Rheumatism Association* 2019:1-8. PMID: 31106666.
48. Narvaez J*, et al.* Predictive factors for induction of remission in patients with active rheumatoid arthritis treated with tocilizumab in clinical practice. *Semin Arthritis Rheum* 2016; 45(4):386-90. PMID: 26254548.
49. Pers YM*, et al.* Predictors of response and remission in a large cohort of rheumatoid arthritis patients treated with tocilizumab in clinical practice. *Rheumatology (Oxford)* 2014; 53(1):76-84. PMID: 24056521.
50. Wang J*, et al.* Relationship Between Baseline and Early Changes in C-Reactive Protein and Interleukin-6 Levels and Clinical Response to Tocilizumab in Rheumatoid Arthritis. *Arthritis Care Res (Hoboken)* 2016; 68(6):882-5. PMID: 26473986.
51. Doorenspleet ME*, et al.* Rheumatoid arthritis synovial tissue harbours dominant B-cell and plasma-cell clones associated with autoreactivity. *Ann Rheum Dis* 2014; 73(4):756-62. PMID: 23606709.
52. Tak PP*, et al.* Dominant B cell receptor clones in peripheral blood predict onset of arthritis in individuals at risk for rheumatoid arthritis. *Ann Rheum Dis* 2017; 76(11):1924-30. PMID: 28790026.
53. Bashford-Rogers RJM*, et al.* Analysis of the B cell receptor repertoire in six immune-mediated diseases. *Nature* 2019; 574(7776):122-6. PMID: 31554970.
54. Adlowitz DG*, et al.* Expansion of Activated Peripheral Blood Memory B Cells in Rheumatoid Arthritis, Impact of B Cell Depletion Therapy, and Biomarkers of Response. *PloS one* 2015; 10(6):e0128269. PMID: 26047509.
55. Mahevas M*, et al.* Long-lived plasma cells in autoimmunity: lessons from B-cell depleting therapy. *Frontiers in immunology* 2013; 4:494. PMID: 24409184.
56. Moura RA*, et al.* B-cell phenotype and IgD-CD27- memory B cells are affected by TNF-inhibitors and tocilizumab treatment in rheumatoid arthritis. *PloS one* 2017; 12(9):e0182927. PMID: 28886017.
57. Pollastro S*, et al.* Non-response to rituximab therapy in rheumatoid arthritis is associated with incomplete disruption of the B cell receptor repertoire. *Ann Rheum Dis* 2019; 78(10):1339-45. PMID: 31217169.
58. Humby F*, et al.* Ectopic lymphoid structures support ongoing production of class-switched autoantibodies in rheumatoid synovium. *PLoS medicine* 2009; 6(1):e1. PMID: 19143467.
59. Yanni G*, et al.* Synovial tissue macrophages and joint erosion in rheumatoid arthritis. *Ann Rheum Dis* 1994; 53(1):39-44. PMID: 8311554.

60. Alivernini S*, et al.* Distinct synovial tissue macrophage subsets regulate inflammation and remission in rheumatoid arthritis. *Nat Med* 2020; 26(8):1295-306. PMID: 32601335.
61. Bresnihan B*, et al.* Synovial tissue sublining CD68 expression is a biomarker of therapeutic response in rheumatoid arthritis clinical trials: consistency across centers. *J Rheumatol* 2009; 36(8):1800-2. PMID: 19671815.
62. Mizoguchi F*, et al.* Functionally distinct disease-associated fibroblast subsets in rheumatoid arthritis. *Nature communications* 2018; 9(1):789. PMID: 29476097.
63. Zhang F*, et al.* Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat Immunol* 2019; 20(7):928-42. PMID: 31061532.
64. Croft AP*, et al.* Distinct fibroblast subsets drive inflammation and damage in arthritis. *Nature* 2019; 570(7760):246-51. PMID: 31142839.
65. Lliso-Ribera G*, et al.* Synovial tissue signatures enhance clinical classification and prognostic/treatment response algorithms in early inflammatory arthritis and predict requirement for subsequent biological therapy: results from the pathobiology of early arthritis cohort (PEAC). *Ann Rheum Dis* 2019; 78(12):1642-52. PMID: 31582377.
66. Lewis MJ*, et al.* Molecular Portraits of Early Rheumatoid Arthritis Identify Clinical and Treatment Response Phenotypes. *Cell reports* 2019; 28(9):2455-70.e5. PMID: 31461658.
67. Humby F*, et al.* Synovial cellular and molecular signatures stratify clinical response to csDMARD therapy and predict radiographic progression in early rheumatoid arthritis patients. *Ann Rheum Dis* 2019; 78(6):761-72. PMID: 30878974.
68. Nerviani A*, et al.* A Pauci-Immune Synovial Pathotype Predicts Inadequate Response to TNFα-Blockade in Rheumatoid Arthritis Patients. *Frontiers in immunology* 2020; 11:845. PMID: 32431716.
69. Dennis G, Jr.*, et al.* Synovial phenotypes in rheumatoid arthritis correlate with response to biologic therapeutics. *Arthritis research & therapy* 2014; 16(2):R90. PMID: 25167216.
70. Kormelink TG*, et al.* Decrease in immunoglobulin free light chains in patients with rheumatoid arthritis upon rituximab (anti-CD20) treatment correlates with decrease in disease activity. *Ann Rheum Dis* 2010; 69(12):2137-44. PMID: 20679475.
71. Raterman HG*, et al.* The interferon type I signature towards prediction of non-response to rituximab in rheumatoid arthritis patients. *Arthritis research & therapy* 2012; 14(2):R95. PMID: 22540992.
72. Vosslamber S*, et al.* Pharmacological induction of interferon type I activity following treatment with rituximab determines clinical response in rheumatoid arthritis. *Ann Rheum Dis* 2011; 70(6):1153-9. PMID: 21444302.
73. Bankhead P*, et al.* QuPath: Open source software for digital pathology image analysis. *Scientific reports* 2017; 7(1):16878. PMID: 29203879.
74. Rivellese F*, et al.* B cell synovitis and clinical phenotypes in rheumatoid arthritis: relationship to disease stages and drug exposure. *Arthritis Rheumatol* 2020; 72(5):714-25. PMID: 31785084.
75. Pagès F*, et al.* International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet* 2018; 391(10135):2128-39. PMID: 29754777.
76. Bashford-Rogers RJ*, et al.* Network properties derived from deep sequencing of human B-cell receptor repertoires delineate B-cell populations. *Genome Res* 2013; 23(11):1874-84. PMID: 23742949.
77. Chaussabel D*, et al.* A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity* 2008; 29(1):150-64. PMID: 18631455.
78. Langfelder P*, et al.* Eigengene networks for studying the relationships between co-expression modules. *BMC systems biology* 2007; 1:54. PMID: 18031580.
79. Bombardieri M*, et al.* Ectopic lymphoid neogenesis in rheumatic autoimmune diseases. *Nat Rev Rheumatol* 2017; 13(3):141-54. PMID: 28202919.
80. Guo Y*, et al.* RNAseqPS: A Web Tool for Estimating Sample Size and Power for RNAseq Experiment. *Cancer Inform* 2014; 13(Suppl 6):1-5. PMID: 25374457.
81. McCarthy DJ*, et al.* Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 2012; 40(10):4288-97. PMID: 22287627.
82. McKeigue P. Sample size requirements for learning to classify with high-dimensional biomarker panels. *Stat Methods Med Res* 2019; 28(3):904-10. PMID: 29179643.
83. Hensor EMA*, et al.* Validity of a two-component imaging-derived disease activity score for improved assessment of synovitis in early rheumatoid arthritis. *Rheumatology (Oxford)* 2019. PMID: 30824919.
84. McKeigue P. Quantifying performance of a diagnostic test as the expected information for discrimination: Relation to the C-statistic. *Stat Methods Med Res* 2019; 28(6):1841-51. PMID: 29978758.