Blood cells phenotyping across the Inflammatory Arthritis Continuum (IAC)

Innocent Chidi Anioke

Submitted in accordance with the requirements for the degree of PhD

The University of Leeds

Leeds Institute of Rheumatic & Musculoskeletal Medicine (LIRMM)

School of Medicine

Faculty of Medicine & Health

December, 2023

Intellectual Property Rights and Publication Statements

I confirm that the work submitted is my own and appropriate credit has been given where reference has been made to the work of others.

A part of this project (at-risk cohort) has been jointly published. I performed most of the data analysis for flow cytometry and statistical modelling while my coauthors provided a clinical overview and contributed to the manuscript editing and revision.

Acknowledgments

I sincerely appreciate my supervisory team (Ass Prof Frederique Ponchel, Prof Savic Sinisa, Prof Paul Emery, and Dr Farag Shuweihdi) for their support, and particularly my led supervisor, Dr. Frederique Ponchel for her relentless support and encouragement towards the successful completion of this work.

I also say a VERY BIG thank YOU to my wife Peculiar, my daughter, Providence, and my son Blossom.

You all have been wonderful and supportive. I love you all

Finally, my praise and all glory belong to the author and finisher of my faith, the LORD JESUS CHRIST for sustaining me till the end of this programme. May HIS name be praised forever and ever. Amen !!

ABSTRACT

Rheumatoid arthritis (RA) manifests through distinct phases along an inflammatory arthritis continuum (IAC), spanning from a pre-clinical phase to diagnosis and treatment. While T-cell subsets have received considerable attention in RA pathogenesis and some of them are used as biomarkers, other lymphocyte subsets (LS) remain under-explored, particularly for their possible value as biomarkers. This study aims to evaluate the biomarker potential value of 18 LS across the IAC, validate a novel dry tube (DT) technology against conventional wet tube (WT) methods, and assess the impact of freezing samples on flow cytometry (FC) analysis.

A total of 210 individuals at-risk for RA comprising progressor (n=93) and non-progressors (n=117), 306 Early Arthritis Clinic evolving to RA (n=206) versus non-RA (n=100), and 205 early RA patients receiving methotrexate (MXT) treatment (106 of whom achieved remission) were selected across the IAC from Leeds observational studies and analysed for 18 LS. Multivariate logistics prediction models were constructed for disease progression, diagnosis, and MXT-induced remission. Results notably revealed enhanced predictive models incorporating B-reg, NK-cells, and CD8+T-cells alongside previously established CD4+T cell subsets for progression in at-risk cohort.

Blood samples (n=41) were tested using DT (designed in collaboration with Becton Dickinson) compared to WT for T-cell and Treg subsets. Using DT significantly reduced turnaround time (about 40%) and error rates however only cell surface staining proved possible and accurate while intra-cellular staining failed to produce data.

The effect of freezing PBMC on cell viability for the enumeration of LC was assessed using matched paired comparison analysis (frozen versus fresh samples, n=15) and the capacity to generate data from 131 frozen samples altogether using 4 FC panels. As expected, freezing PBMC greatly reduced cell viability compared to fresh samples, with a threshold of ≥15% viability enabling accurate LS enumeration. Furthermore, freezing altered the ability to detect certain T-cell subsets due to the sensitivity of some markers (notably CD62L) to freezing.

This study provides several novel insights. First, with respect to the pathogenesis of RA, new clues as to events that could be involved in the progression to RA in the pre-clinical disease course were identified while my work offered additional stratification markers that could be used stratification towards clinical prevention studies. Secondly, it demonstrated the feasibility of using DT for T-cell subset enumeration with a gain of time and capacity for standardisation of this assays. Lastly defined a rule (viability limit & need for adapting markers) that would ensure reliable FC data can be obtained from frozen samples allowing for multicentre/retrospective studies where freezing is inevitable.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Abbreviations

RA	Rheumatoid arthritis
OA	Osteoarthritis
AS	Ankylosing spondylarthritis
PsA	Psoriatic arthritis
ACPA	Anti-citrullinated protein antibody
RF	Rheumatoid factor
IAC	Inflammatory arthritis continuum
BS	B-cell subsets
LS	Lymphocytes subsets or subpopulations
SOP	Standard operation procedure
FCS	Flow cytometry standardise (file format)
CRP	C-reactive protein
DAS	Disease activity score
PPV	Positive predictive value
NPV	Negative predictive value
AUROC	Area Under the Receiver Operating Curve
CD	Cluster of differentiation
EAC	Early arthritis clinic
ESR	Erythrocyte sedimentation rate
EMS	Early morning stiffness
HC	Health control
MXT	Methotrexate
RXT	Rituximab
TJC	Tender joint count
SJC	Swollen joint count
SE	Shared epitope
HLA	Human leukocyte antigen
FSC	Forward scatter
SSC	Side scatter
PPC	Pre plasma cell
Breg	Regulatory B-cell
Treg	Regulatory T-cell
NK	Natural killer cell
DMARD	Disease modifying anti-rheumatic drugs
UA	Undifferentiated arthritis
CTLA4-I	cytotoxic-T-lymphocyte-associated 4–immunoglobulin)

APCs	Antigen presenting cells
DCs	Dendritic cells
RANKL	Receptor activator of nuclear factor κB ligand
TNF-α	Tissue necrotic factor- alpha
IL	Interleukins
EULAR	European League against Rheumatism
IA	Inflammatory Arthritis
NICE	National Institute for Health and Care Excellence
IRC	Inflammatory related cells
PBMC	Peripheral blood mononuclear cell
ERA	Early rheumatoid arthritis

Chapter 1
Introduction

1.1 Biomarker translational research

The foundation for precision medicine begins with better patient stratification and accurate prediction of heterogeneous outcomes associated with most disease conditions. Identification of robust predictive biomarkers will be important to lay such a foundation. In a disease model such as rheumatoid arthritis (RA), this was shown to be achievable with a biomarker detection technology such as flow cytometry (FC) [1-9]. Biomarkers, besides enabling the identification of individuals who are likely to progress to a particular outcome, have the greatest potential for guiding informed clinical decisions. Furthermore, biomarkers being rooted in the pathophysiologic mechanism of diseases and reflecting the disturbances of cellular pathways directly involved in their pathogenesis are more likely to be robust [10, 11].

In RA, immune cells are likely to offer such biomarkers value as deeply related to the pathology hence being able to measure their characteristics is therefore highly desirable.

This PhD will investigate the value of immune cells as biomarkers for the management of RA. The technology of choice for biomarker measurement here will be FC.

1.2 Rheumatoid Arthritis (RA): A disease model for validation of FC biomarkers

RA is the most prominent member of a vast number of autoimmune-mediated inflammatory diseases with a significant burden on patients and society [12]. It is also a form of chronic systemic inflammatory polyarthritis with an autoimmune component. It is characterised by swollen, stiff, and progressive destruction and deformity of the joints of the hands, knee, and feet [13, 14] occurring very rapidly in 80-85% of patients within the 1st 2 years of the disease onset [15, 16]. It affects ~0.5% to 1% of the overall western population [17], 0.06 to 3.4% across Africa and the Middle East [18] and more females than males at a ratio of 3:1 [19, 20]. The disease affects greater than >450,000 of the UK population with a 32% increase in the risk of mortality compared to people of the same age without the disease [21-24] with about 1.5 men and 3.6 women developing RA per 10,000 people per year in the UK [25]. Furthermore, the functional capacity in RA patients gradually deteriorates with time [26].

1.2.1 Clinical signs and symptoms of RA

The characteristic features of the disease are symmetric polyarthritis affecting the small joints of the hands, knee, and feet, with early morning stiffness (EMS) [27]. Despite predominant articular manifestations, the disease progression may vary from mild, or self-limiting in some cases [28, 29] to an aggressive form [30] that may involve extra-articular manifestation [31]. With approximately 1% of the UK population suffering from RA, and ~15% of patients having severe disability impaired quality of life, and reduced life expectancy [13, 32, 33]. RA is a serious, painful, and debilitating condition, despite not being immediately life-threatening.

1.2.2 The RA disease burden: economic cost and personal burden

The incidence and prevalence of RA are increasing with considerable morbidity worldwide [12, 34-40]. Although treatment strategies and management approaches have improved substantially, only a few patients (about 20%) nowadays achieve and sustain clinical remission for those periods of time. However, accessing RA treatment is undoubtedly a challenge complicated by the high cost of certain drugs (ie biologics) on limited healthcare budgets [41]. Consequently, this high cost of treatment, creates psycho-social and economic burdens for the patients, affecting their quality-of-life [42-47]. With RA affecting individuals in their productive age (mid 40 – 60 years) with the associated functional disabilities, it is conceivable that the indirect cost of RA may outweigh direct patient care.

Costs relating to disabilities and reduction in earning capability as well as reduced life expectancies are, often difficult to quantify and constitute the indirect cost of RA also termed intangible cost [48-50]. Many RA patients (up to 90%) also experienced chronic fatigue with a significant level of absenteeism from work due to fatigue [51, 52], or not being able to work leading to about 71% job loss for patients at working-age [53, 54], while the Early RA Network inception cohort study (ERAN) reported that 10% of patients even with the early RA stage could not sustain their jobs. Similar findings have been reported in the USA [55, 56]. The greater risk of disability among RA patients has been linked to depression [57, 58] and persistent fatigue [59] contributing further to the reduced work ability experienced by RA patients.

Direct costs representing patient management, include diagnostic, hospitalization, therapeutic intervention, mechanical aid for patients, physiotherapy, and nursing time [41]. It varies between countries given the variation of healthcare organisations) [48, 60] It is evident that RA is associated with functional and psychological disabilities that pose an added substantial healthcare burden and costs [22] including, but not limited to, hospitalizations in up to 8.6%, of patients in Germany while a bit less in the UK or Spain [61]. RA flare (35%) and pulmonary infections (25%) are top on list for RA-related hospitalisation, confirming the high healthcare resource utilization from RA management. This also includes the frequency of occupational and physiotherapies used by patients, with a greater percentage of UK patients using both services (30.1%/44.7%) compared to patients in Germany (11.8%/33.3%) or Spain (2.9%/15.5%)[61].

There is also increased functional damage in the late disease stage, necessitating the need for orthopaedic surgery in most patients [62]. This highlights another huge direct cost and impact on the patients' economy adding to the burden of rheumatic disease (such as OA) [63]. Evidence has shown also that those who did not get access to effective treatment early in their disease were hospitalized more often than those treated early [61].

Later in the disease course, comorbidities such as infection (due to treatment related immune suppression), depression, and cardiovascular disease are a great concern and source of worry for RA patients [64, 65].For instance, RA patients have an increased risk of developing cardiovascular disease compared to the general population. This increased risk is attributed to chronic inflammation, traditional cardiovascular risk factors, and the use of certain RA medications [66].

A study emphasized the importance of addressing both RA disease activity and cardiovascular risk factors to optimize treatment outcomes and reduce the burden of comorbidities in RA patients [67]. The burden of the disease is far greater in patients with severe disease than in those with low or mild disease activity [68]. This suggests the need for an effective treatment approach to achieve a low disease activity that would reduce patients' pain, burdens as well as hospitalizations, and its associated costs due to accumulation of damages as well as added burden of developing comorbidities.

In addition to the holistic approach of RA management, targeting depression as part of routine clinical care could be a positive option towards reducing the likely economic costs for the patient, family, and society [22, 69]. **Table 1** summarises the significant contribution of RA on the economic burden both of patients as well the society [70] particularly in Western countries where work-related disabilities and sick leaves cost the economy some billions of dollars [71].

Although RA management seems to have improved substantially over the years, the huge economic burden still associated with RA and impacting not only the patients but also the society at large highlights gaps in the RA management approach and the need for reducing, if not reversing, the debilitating consequence of RA in individuals of productive age.

Table 1 Country estimated cost for RA management.

Country	Direct and indirect cost	Ref
UK	£3.8 and £4.8 billion (US$5.1–6.4 billion)[	[23, 69]
USA	~ $128 billion	[72]
Canada,	$4.4 billion	[73]
Europe;	~ €3–5000 per patient (US$3589-5982)	[69]
• **France**	222 million Euros	[74]

1.2.3 Established risk factors for RA

Established risk factors for RA encompass a multifaceted array of genetic, environmental, and lifestyle influences. An understanding these risk factors would not only aid in early detection but also informs preventive strategies and therapeutic interventions, thereby shaping approaches to mitigate the impact of this chronic autoimmune disease.

1.2.3.1 Genetic risk factor

There is an established role for genetic factors in the risk, progression, and severity of RA [75]. First-degree relatives remain on top of the established risk factors conferring up to 2-4-fold increased risk for developing RA [76]. This has therefore confirmed the genetic contribution to RA susceptibility, also verified in twin studies suggesting an overall 30% genetic contribution [77, 78]. The strongest genetic association with RA is located on the (Human leukocyte antigen – disease receptor Beta 1 (HLA–DRB1)- locus, in particular a specific amino-acid motif in the gene (the “shared epitope" (SE) which has been linked to RA susceptibility and estimated to account for >65% of heritability [79, 80], while leaving other risk factors account for the remaining 35-45% variance in

developing RA [81]. Other studies have, however, reported much lower estimated heritability (~50% for anti-citrullinated protein antibodies (ACPA)-positive RA and ~20% ACPA-negative), suggesting a lower genetic risk contribution of SE [80, 82-84]. Several assumptions and the different methods used in heritability estimation by different studies [76] may perhaps explain these inconsistencies in results. RA severity, mortality and treatment response have also been linked to different haplotypes of HLA-DRB1[85].

Genome-wide studies have further associated multiple genetic loci, defining >100 non-HLA risk loci [79, 86, 87]. The majority of these genes are related to T-cell immune mechanisms and antigen presentation [8, 88] such as Protein tyrosine phosphatase non-receptor type 22 (PTPN22), cytotoxic T-lymphocyte antigen-4 (CTLA4) and Signal transducer and activator of transcription 4 (STAT4), which are implicated in T-cell activation and might account for about 5% variance in genetic risk associated with developing RA [79, 89]. PTPN22 has been associated with an increased risk of developing autoimmune diseases, including RA. This gene encodes a phosphatase that negatively regulates T-cell receptor signalling which may lead to dysregulated immune responses and contribute to the development of the disease [90, 91]. CTLA4 is a protein receptor that downregulates immune responses by inhibiting T-cell activation. Genetic variations in CTLA4 have been linked to susceptibility RA. Reduced expression or function of CTLA4 may result in unchecked T-cell activation and contribute to autoimmune inflammation seen in RA [92, 93]. STAT4 is a transcription factor involved in the signalling pathways of several cytokines, including interleukin-12 (IL-12) and interferon-alpha (IFN-α). These cytokines play crucial roles in the differentiation and activation of T-cells and other immune cells. Dysregulated STAT4 signalling may contribute to aberrant immune responses and

inflammation observed in RA patients [94, 95]. Furthermore, single-nucleotide polymorphisms (SNPs) coding for molecules that regulate T-cell activation have also been associated with RA [86]. However, these loci appear less valuable in predicting RA development due to their low effect size [96].

Altogether, only about 13-18% of the total RA genetic risk has been accounted so far by all currently known genetic risk factors combined [79, 86], which suggests that genetic predisposition is unlikely the sole determinant of heritability. Evidence of selected gene–environment interactions in RA development has been proposed, notably for smoking and SE [80, 96, 97], indicating that such factors may likely play a ma or contributory role.

1.2.3.2 Environment risk factors

There has been a substantial interest in associating environmental and lifestyle risk factors with the pathogenesis of RA. A considerable amount of evidence has demonstrated a close interaction between the gene loci-HLA-SE and smoking [98, 99]. Besides smoking, other factors such as infectious organisms [100, 101], air pollutants (fine particulate (PM 2.5) [102, 103], ozone exposure, and living near high-traffic roads [104, 105] have also been implicated in the risk of developing RA. Most results come from single-centre studies and results across multiple studies have also been inconsistent. Therefore, only smoking has shown a relatively consistent association with RA as the strongest established environmental risk factor estimated to account for about 20% to 30% of this risk [106].

1.2.3.3 Genetic-Environmental interaction

Gene-environment studies showed that positivity for HLA-DRB1 alleles (SE) and the PTPN22 polymorphism constitute the major risk resulting in the production of rheumatoid factor (RF) and ACPA with smoking [107, 108]. While the driving factors responsible for autoantigen generation (citrullinated self-protein) which serves as the trigger for autoantibodies production remain unclear, a range of inflammatory processes such as non-specific local inflammation and formation of neutrophil extracellular traps (NETs) as a result of gingivitis, bad mouth hygiene and dysbiosis of the buccal microbiota may be implicated [109-113]. NETs are web-like structures composed of DNA, histones, and antimicrobial proteins released by neutrophils to trap and kill pathogens. NETs have been implicated in promoting inflammation, autoimmunity, and tissue damage. They can activate immune cells, stimulate the production of pro-inflammatory cytokines, and contribute to the formation of autoantibodies [113-117]. For instance, NETs contain citrullinated autoantigens, such as citrullinated histones and other proteins. These citrullinated proteins are recognized by the immune system as foreign antigens, leading to the production of autoantibodies, such as ACPAs [113]. ACPAs are characteristic markers of RA and contribute to the autoimmune response in RA patients. Additionally, they release pro-inflammatory molecules, including cytokines and chemokines, which activate immune cells such as macrophages and dendritic cells. These activated immune cells further propagate inflammation within the joints, leading to synovial inflammation and damage to cartilage and bone [115, 116].

There are several mechanisms proposed to mediate this gene-smoking interaction. One mechanism suggests that in the context of HLA-SE, smoking plays a role through the citrullination of self-protein and subsequent production of

autoantibodies [118]. For instance, increasing levels of expression of extracellular peptidyl arginine deiminase 2 (PAD2) in the lungs plays a role in the citrullination of mucosal proteins [98]. Smoking was also shown to alter the DNA methylation pattern of the *HLA-DRB1*, potentially affecting it expression [119, 120].

Another mechanism proposes a role for local inflammation at mucosa surfaces [120-122] such as periodontal and pulmonary disease, also observed in smokers [100, 101]. It has been suggested that biological interactions occurring at these sites might promote inflammation and autoimmunity but whether these mucosa surfaces are the initial site of autoantibodies development is not fully understood [123-125]. Smoking has been associated with elevated bacterial colonisation as well as microbiome dysbiosis of the pulmonary mucosal surface [126], and some studies have provided evidence of an association between such dysbiosis and an early breach in tolerance mechanisms [127-129].

Microbial mucosa-restricted processes influencing RA pathogenesis have implicated a few bacteria, particularly those involved in periodontal infection given their specific role in initiating citrullination of host proteins leading to the generation of RA-related autoantibodies. Infectious organisms such as *Porphyromonas gingivalis* [100, 101], Anaeroglobus and Prevotella species [130] as well as *Aggregatibacter actinomycetemcomitans* [131] responsible for periodontitis, have been positively associated with autoimmunity in RA. The model underpinning their association with RA involves the production of enolase and bacterial *(Porphyromonas gingivalis)* encoded Peptidylarginine Deiminase (PAD) leading to citrullination of self-protein or production of leukotoxin-A (*Aggregatibacter actinomycetemcomitans)* inducing hypercitrullination in human protein (notably presented on NETs) which potentially leads to generation of autoantigens and trigger ACPA production.

Porphyromonas gingivalis a key driver of periodontal inflammation has been identified as responsible for PAD-mediated citrullination of the mucosa surfaces and may play a critical role in the initial break of immune tolerance [132-134]. Increased expression of human PAD-2 and PAD-4 (protein and mRNA) from the local periodontal tissue as well as citrullinated proteins in inflamed gingiva both support the hypothesis that periodontal inflammation contributes to the inflammatory burden that may lead to the breach in tolerance [135]. The importance of the site of antigen presentation, such as local periodontal tissue, in RA development lies in the concept of "molecular mimicry" and the initiation of autoimmunity [136-138]. Periodontal pathogens, such as *Porphyromonas gingivalis* possess proteins with structural similarities to host proteins, including citrullinated proteins. When the immune system responds to the presence of these pathogens in the periodontal tissue, it may also produce antibodies that cross-react with similar epitopes on host protein and can lead to the development of autoimmunity [100, 139-142].

1.2.3.4 Comorbidities as a risk factor for RA

Pre-existing comorbidities may also impact disease onset and progression. Conditions such as cardiovascular diseases (CVDs) [143], osteoporosis [144], and certain infections [145] discussed in the previous section have been linked to heightened RA risk [64-66]. Disorders like lupus and psoriasis often coexist with RA, suggesting shared underlying mechanisms [146]. Metabolic syndromes, including obesity and diabetes, also contribute to increased susceptibility [147]. Comorbidities represent important risk factors for RA development and contribute to disease severity and treatment outcomes. Addressing comorbidities as part of

comprehensive RA management is crucial for improving overall health outcomes and reducing the burden of disease in RA patients.

1.3 Pathogenesis of RA

The exact pathogenesis of RA is not yet fully understood. However, a model describing the aetiological hypothesis of RA proposed that specific environmental factors such as smoking altogether in a genetic-restricted context (HLA-DR shared epitope), may trigger both cellular and humoral immune response with the potential to contribute to RA pathogenesis [120] as illustrated in **Figure 1.** RA is, however, a very heterogeneous disease with various clinical presentations. Some autoimmune processes are thought to be involved in addition to infiltration of immune cells into the joint tissue resulting in inflammation, cell alterations, expression of cytokine, and other mediators, notably enzymes of tissue degradation [14, 75, 148-151] .

The autoimmune processes are known to play a role as evidenced by linkage with HLA-SE [152, 153], autoantibodies production (anti-citrullinated protein antibody, ACPA and rheumatoid factor, RF) [154-159]. Both features supported the hypothesis of a T-cell-related disease, developed in the late 80s [160-162] and backed up more recently by the genetic risk associated with many T-cell-related genes [8, 88, 163]. The successful response to a therapeutic intervention targeting T-cells with abatacept [163, 164], further strengthens this hypothesis. However, there is very little T-cell clonality in RA and cytokines are expressed by many other cells [165, 166]. The T-cell centric hypothesis has therefore always been challenged while never disproven.

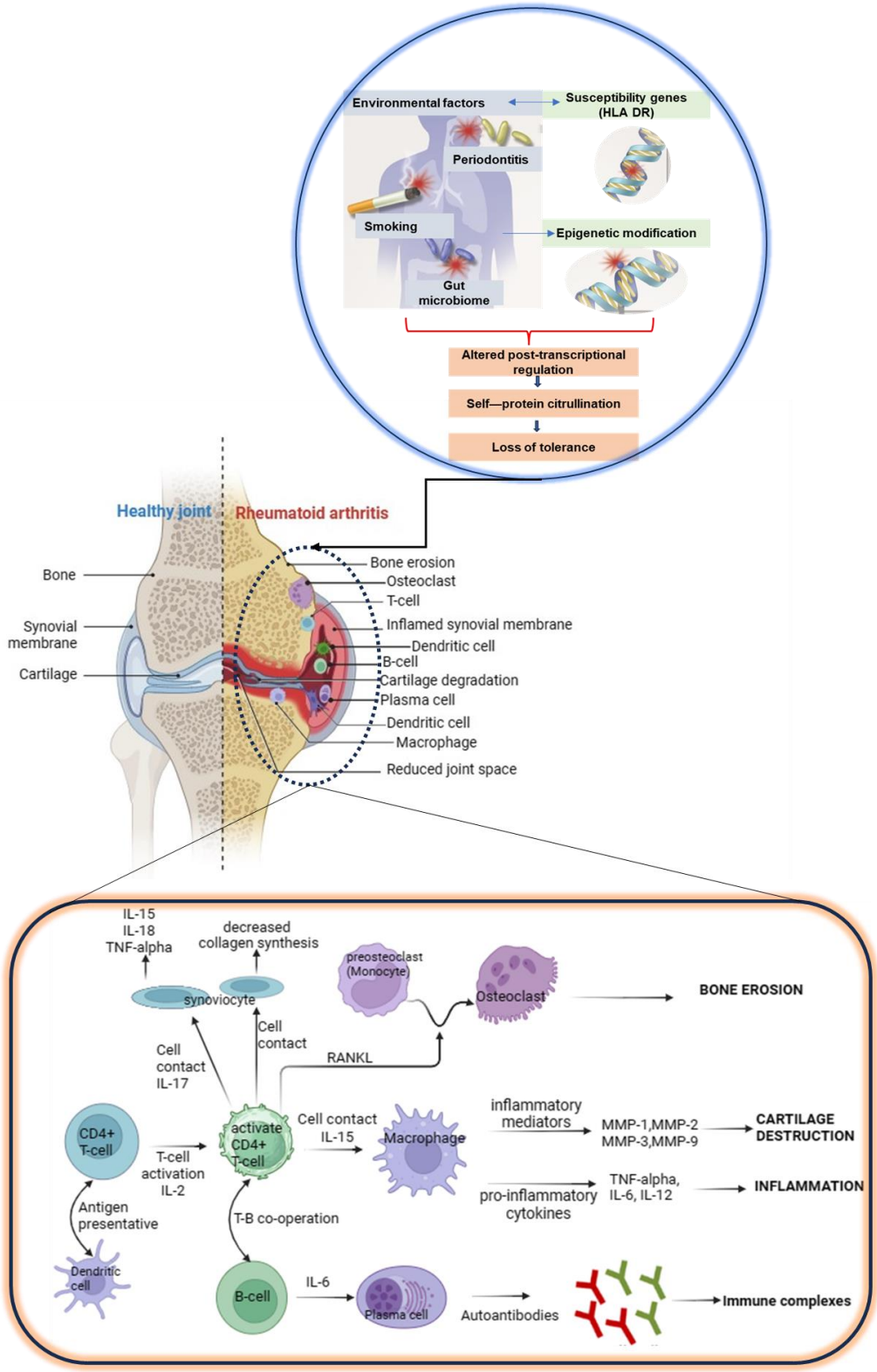


Figure 1: An overview of RA pathogenesis.

The figure illustrated the key players driving RA pathologic events ranging from gene-environment triggers and immunological signatures leading to clinical synovitis and bone damage. Receptor Activator of Nuclear factor Kappa B Ligand (RANKL), interleukin (IL), Tumour necrosis factor (TNF), Matrix metalloproteinases (MMPs). Figure Created with BioRender.com**.**

1.3.1 Systemic autoimmunity as an early biomarker of the breach of tolerance/autoimmunity

Before the onset of clinically evident inflammatory arthritis (which is presented on a physical examination as a swollen joint), there is a phase of detectable systemic autoimmunity. This is a period when RA-related autoantibodies, notably ACPA and/ or RF are detectable [167-175]. However, the key biological events that lead to the initial breach of self-tolerance at this early stage of the disease are yet to be fully elucidated. While not consistent across all studies [176], several risk factors have been identified that could play a role in the initiation, and propagation of autoimmunity before the development of clinically evident inflammatory arthritis. This early stage, RA is also characterised by autoantibodies response to a small number of self-antigens [172-175] and the development of systemic inflammation [177, 178].

Experimental arthritis model of the initial events leading to a breach of self-tolerance showed that a Th1 and Th17 response to antigen resulted in arthropathy [179]. This was characterised by IFN-γ production (rather than lL-17) and associated with spontaneous induction of autoreactive T and B cell response. These early events are thought to be followed by the expansion of autoimmunity and systemic inflammation over time, with the development of innate [180] and adaptive responses [3] , increased widening of the ACPA isotypes or repertoire (epitope spreading) and other autoantibodies, and subclinical inflammation before progression to IA occurs [169]. ACPAs belong to different immunoglobulin isotypes, including IgG, IgM, and IgA. However, IgG ACPAs are the most studied and clinically relevant isotype in rheumatoid arthritis (RA) due to their strong association with disease severity and diagnostic utility [181, 182] .

In humans, the presence of autoantibodies directed against citrullinated and carbamylated antigens before clinical synovitis manifests demonstrated that breakdown in immune tolerance precedes RA manifestation [183], particularly in genetically predisposed individuals. Evidence of the absence of synovitis in the presence of elevated autoantibodies (ACPA/RF) in the peripheral circulation further strengthens this hypothesis [172-175, 184, 185].

The primary site where this event is initiated remains controversial. However, relevant events leading to the initial systemic breakdown of tolerance are thought to start outside the joint [186]. A defect in T-cell peripheral tolerance is believed to contribute, whereby T-cells specific for post-translational modified antigens escape thymic selection [186]. This results in a thymic education that does not completely remove T-cells recognising such protein modifications. The fact that, RA-specific autoantibodies are not directed against self-peptide sequences but against a protein that has undergone the same type of specific post-translational modifications (PTM) results in a high degree of multi-reactivity [187, 188]. Furthermore, loss of tolerance to PTM self-peptide is characterised by epitope spreading leading to the expansion of the repertoire of peptides recognised notably by ACPAs as the disease progresses [189]. Avidity maturation of ACPAs increases from the phase of manifestation of systemic autoimmunity towards the time when RA diagnosis is made [189, 190] suggesting that the maturation of the autoimmune responses directed against citrullinated self-antigens may be associated with disease progression across the IA.

1.3.2 Characterisation of Cell and cytokines interaction (cell and pathways)

Antigen-presenting cells (B-cell, macrophages) via human leukocyte antigen (HLA) or major histocompatibility class II (MHC-II) molecules are efficient in presenting citrullinated self-proteins (autoantigens) to CD4+ T-cells [107, 191, 192]. The interaction of these T-cells with autoreactive B-cells arising from impaired central and peripheral B-cell tolerance checkpoints in RA patients could subsequently lead to the production of autoantibodies (e.g., ACPA) in response to a large spectrum of citrullinated proteins [120, 191, 193-198]. However, T-cell frequencies for such post-translational modified (PTM) proteins are not high if detectable at all [186].

T-cell

T-cells are effectors of cellular immune response which mature in the thymus and expand rapidly upon the encounter of their cognate antigen under the influence of IL-2 [199-201]. They function as helper-CD4+ T-cells or cytotoxic CD8+T-cells, providing cytokine signals to enhance B-cell responses through direct recognition of specific antigens [199].

The role of T-cells in RA pathogenesis has been the subject of numerous studies, recently reviewed in [202]. In the last ten years, extensive work has been done to understand the function of T-cells in RA. Even though the pathogenic role of T-cells is poorly understood, evidence indicates that T-cells are actively involved in the persistence of inflammatory immune responses in RA [203].

For a long time, it was believed that Th1 cell secretion of proinflammatory cytokines was the primary factor in the pathogenesis of RA. CD4+ T-helper cells, notably Th1 cells may be important actors in the pathogenesis of the disease. In RA patients, they produce and secrete a wide range of essential inflammatory mediators, such as tumour necrosis factor (TNF), interferon-gamma (IFN), and IL-2 which are important players in cell-mediated immunity and can activate macrophages into antigen-presenting cells (APCs), which enhances the antigen presentation process via MHC-II molecules [202, 204]. However, their activity in RA was shown to be defective being biased to IFN-gamma cytokine production only rather than all [205-207] as well as not completely polarised (Th1 expression not being increased as much) and not fully activated with IL-2 production notably lower [208]. On the other hand, Th2 polarisation and capacity appear to be untouched in RA [209].

The initiation and maintenance of chronic inflammation are mediated by a variety of T-cell phenotypes and their associated effector mechanisms (**Figure 2**)**,** which contribute to the development of RA [210]. The identification of additional CD4+ T-cell subsets, such as Th17 and Treg [211-217], Th2 [218], Th9 [219, 220], and Tfh and Tph cells [220, 221], however, have provided more recent insight showing the involvement of the T-cells in RA pathogenesis.

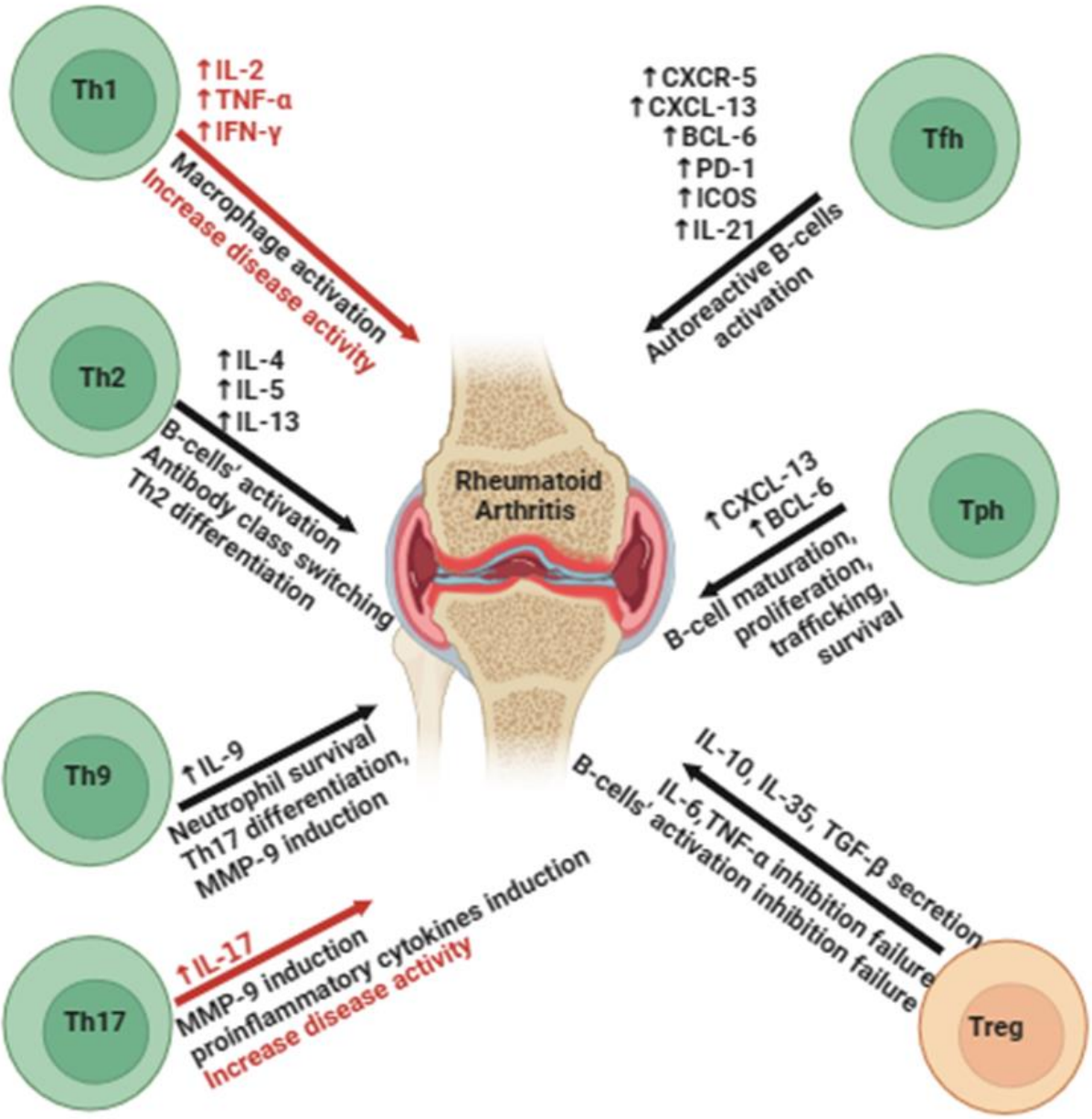


Figure 2 : The role of T-cell subsets in RA.

Briefly, in patients with rheumatoid arthritis (RA);

- T-helper 1 (Th1) cells exhibit heightened activity, releasing pro-inflammatory mediators like tumour necrosis factor-α (TNF-α), Interleukin-2 (IL-2), and interferon-gamma (IFN-γ), fuelling disease activity [202, 221]
- T-helper 2 (Th2) cells secrete cytokines such as IL-4, IL-5, and IL-13. These anti-inflammatory cytokines are primarily involved in activating B cells, facilitating antibody class switching, and promoting Th2 differentiation [218]
- T-helper 9 (Th9) cells are increased in the synovial fluid and tissues of RA patients. IL-9 derived from Th9 cells may play a role by promoting Th17 differentiation, MMP-9 production, and enhancing neutrophil survival [219, 220]
- T-helper 17 (Th17) cells secrete IL-17, which stimulates the production of matrix metalloproteinases (MMPs), chemokines, and pro-inflammatory cytokines which fuel disease activity [211, 212]

- T follicular helper (Tfh) cells can engage with B cells, activating autoreactive B-cells and initiating autoantibody production. They express elevated levels of CXCR5, B-cell CLL/lymphoma 6 (BCL6), C-X-C motif chemokine ligand 13 (CXCL13), the inducible T cell co-stimulator (ICOS), programmed death-1 (PD-1), and IL-21, all crucial for interacting with B-cells [222, 223]
- T peripheral helper (Tph) cells, characterized by low levels of B-cell lymphoma 6 protein (BCL6) and lack of C-X-C chemokine receptor type 5 (CXCR-5) expression in RA synovial tissue, significantly influence B-cell maturation, proliferation, trafficking, and survival upon activation [222]
- Tregs primarily suppress autoimmune processes by secreting anti-inflammatory cytokines like IL-10, IL-35, and transforming growth factor-beta (TGF-β). However, while Treg cells can restrain the proliferation of effector T-cells, they are unable to inhibit the production of inflammatory mediators such as IL-6 and TNF-α as well as restrain B-cell activation [213-217]

Our group has documented a significant alteration in T-cell dynamics in RA, demonstrating that inflammatory processes drive abnormal T-cell differentiation in RA [7]. Essentially, T-cell differentiation following the initial activation of an antigen-inexperienced "naive" T-cell results in the development of a "memory" phenotype, enabling a quick and forceful response to a secondary challenge, that lasts for many years after the initial antigenic encounter. It involves numerous molecular processes that are still poorly understood but which transform naive cells into highly responsive cells although very restricted, quickly dividing into memory cells.

While CD8+ T-cell differentiation has been extensively studied, CD4+ T-cell differentiation appears to differ in several ways. Naive CD4+T-cell (Tho) activation leads to a commitment to either a Th1 or a Th2 phenotype (or other polarised cells) [202]. This entails the epigenetic reprogramming of a certain set of genes and the commitment of a particular set of genes to produce either Th1 and other polarisation .

Various theories (with a focus on CD8+ T-cells) have been proposed over time to explain the transition from naive to memory cells. In the past, a linear model proposed **(**Error! Reference source not found.**)** that an effector phenotype d evelops directly from naive cells, resulting in most of the effector population dying and a small number of quiescent memory cells surviving at the end of a response [224].

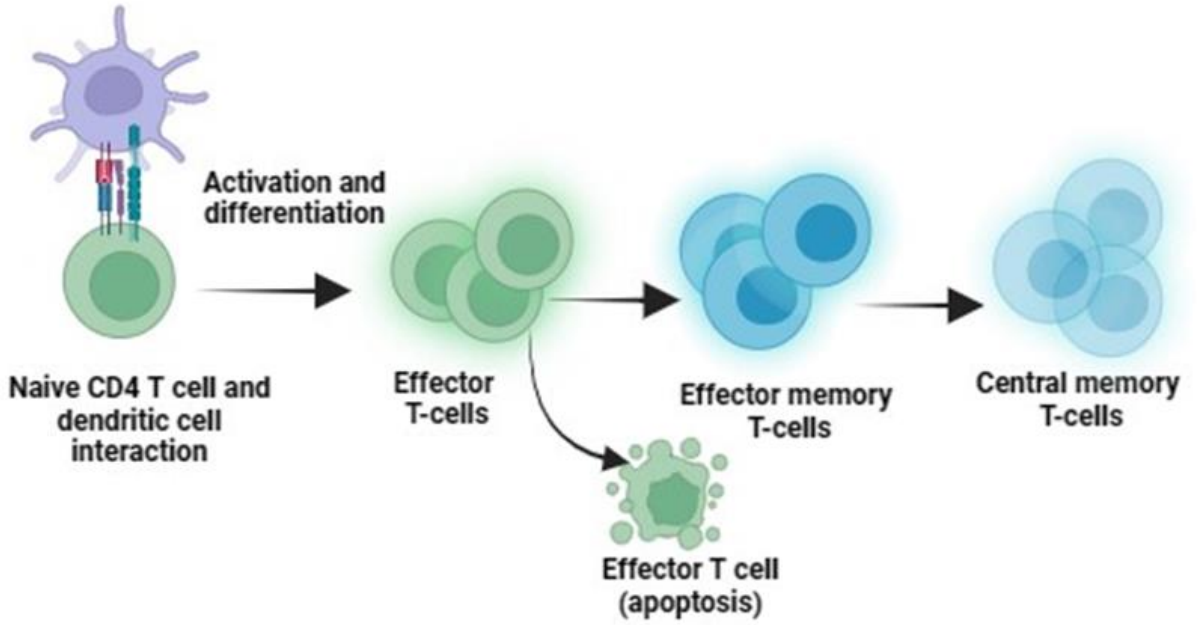


Figure 3 Linear model of T cell differentiation**.**

Model illustrate transition from naïve subset to memory subset upon initial activation. Figure created in BioRender.com

There are now other models emerging. One model (illustrated in Error! R eference source not found.) suggests a non-linear differentiation were naive cells can be turned into effector cells and then memory cells, but the point at which differentiation stops depends on the strength of the initial signal [225, 226]. Th0 naïve cells will transform into a memory phenotype, if the signal from the local environment is strong (IL-2), with either a Th1 or Th2 drive if present, and finally if there is antigen persistence. Alternatively, Th0 naïve cells will adopt a non-effector phenotype and develop into memory-like cells with a preference for lymph nodes where they can produce secondary responses if the signal is weak (low IL-2) in the presence of IL-15 or transforming growth factor-beta (TGF-β).

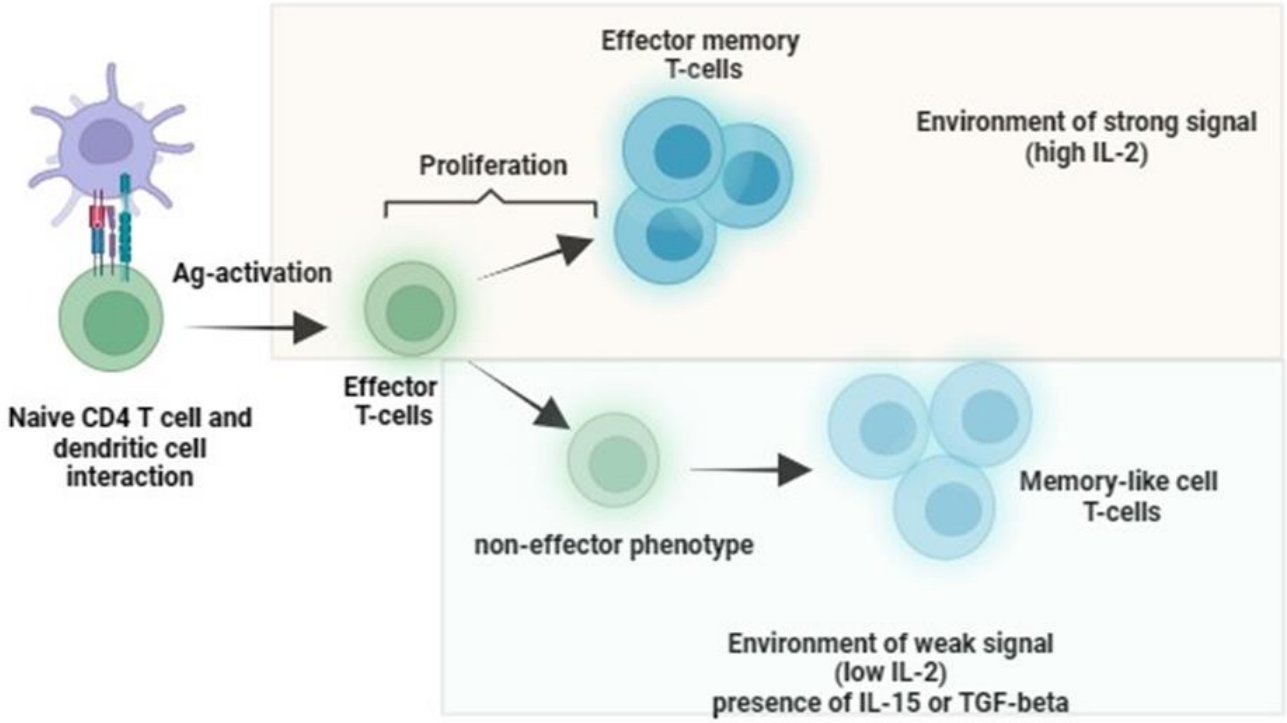


Figure 4 Alternative model of T-cell differentiation.

Model illustrate transition from naïve subset to memory subset depending on strength of inflammatory signals. Figure created in BioRender.com

Data underpinning this model remain inconsistent regarding the precise phenotype of these memory cells and their longevity over time. Proliferating cells also quickly experience clonal senescence(irreversible growth arrest or ageing), making it impossible to maintain them indefinitely through repeated proliferative cycles. Highly specialized memory T-cells need survival factors (for example IL-7) to be maintained because they are more susceptible to apoptosis [224]. Cellular senescence is a state of irreversible growth arrest which a group of cells derived from a single progenitor cell (a clone) undergoes simultaneously or sequentially due to shared stress signals or intrinsic factors, including telomere shortening, DNA damage, and oncogene activation [227] . However, a few studies claim that memory CD4+ T-cells can exist in a quiescent state with a phenotype similar to that of naive cells (CD45RA) exhibiting lower apoptosis

susceptibility and therefore being capable of long-term memory, and called central memory cells [228, 229].

Our team, the Translational Research in Immune-Mediated Inflammatory Disease (TRIMID), has examined T-cell differentiation in a cross-section of healthy individuals and RA patients, paying special attention to CD4+T-cells in the early stages of the disease and correlating it with an examination of the activity of the thymus. T-cell precursors rearrange their T-cell receptor genes as they move through maturation releasing tiny circles of episomal DNA known as T-cell receptor excision circles (TRECs) [230, 231]. Therefore, an estimation of thymic function can be made by counting the percentage of peripheral T-cells that contain TRECs. However, many dynamic processes, particularly peripheral T-cell proliferation, have an impact on TREC measurements [7, 232, 233]. So the number of cell cycles that take place prior to reaching that differentiation stage is reflected in the TREC content of a given sub-population of cells. With this understanding, our team has put forth a model of T-cell differentiation in healthy individuals by combining differentiation markers and TREC measurements [7].

The two models previously discussed both agree and disagree with the T-cell differentiation model in some ways, but direct comparison between the three models is challenging due to the different markers that were employed.

Despite RA pathogenesis being considered a T-cell-mediated process, antigen-driven clonal expansion of T-cells has not been demonstrated in RA [234]. Conversely, interaction (i.e., cell-cell contact) with antigen-presenting cells (APCs) (macrophages, dendritic cells (DCs), and B-cell), as well as fibroblasts, was observed in the synovium [235, 236] , suggesting another possible role for T cells as the orchestrator of a cell-cell interaction network [235, 237]. T-cells

triggering B-cell responses and antibody production maturation may also ultimately contribute to the pathogenesis of RA [238].

B-cell

B-lymphocytes are immune cells primarily mediating adaptive humoral immunity via an antibody-dependent mechanisms. B-cells also function as APCs and produce both anti- and pro-inflammatory cytokines [199]. In the circulation, several B-cell subsets **(Figure 5)** can be recognized based on cell surface-marker expression, which represents their differentiation status [199, 239]. The involvement of circulating and tissue-localized B-cell subsets in RA can manifest through various mechanisms, such as autoantibody generation via antigen presentation, T-cell stimulation, and cytokine production [10, 240, 241].

Evidence for a role of B-cells in autoimmunity has been described independently of autoantibody production in RA, via the production of and response to pro-inflammatory cytokines, and/or co-stimulatory activation of T-cells [240-243], and notably due to the success of rituximab, a B-cell depletion therapy [12, 244] which depletes B cells but does not affect antibody production and only marginally reduces ACPA level [245, 246] probably by depleting short-lived plasmablasts but not long-lived ones (as they do not express CD20) [247].

The pathogenesis of RA has been associated with B-cell functions. One popular hypothesis of how B-cells can affect the course of an inflammatory response is the production of autoantibodies (RF and ACPA) by autoreactive B-cells after differentiation into plasma cells [248-252]. A second model proposes that B-cell involvement in RA pathogenesis is by cytokines production. TNF, IL-10, and IL-6, [253, 254] as well as the immunoregulatory cytokines IL-2, INF-γ, IL-12, and

IL-4, can also be produced by B-cells in RA [255]. Particularly, it has been suggested that the cytokines TNF, IL-6, and INF-γ play a significant role in RA development. A third hypothesis suggest that antigen presentation and activation of antigen-specific T helper cells is another critical component of B-cell function that plays a role in the onset and progression of RA [256-260].

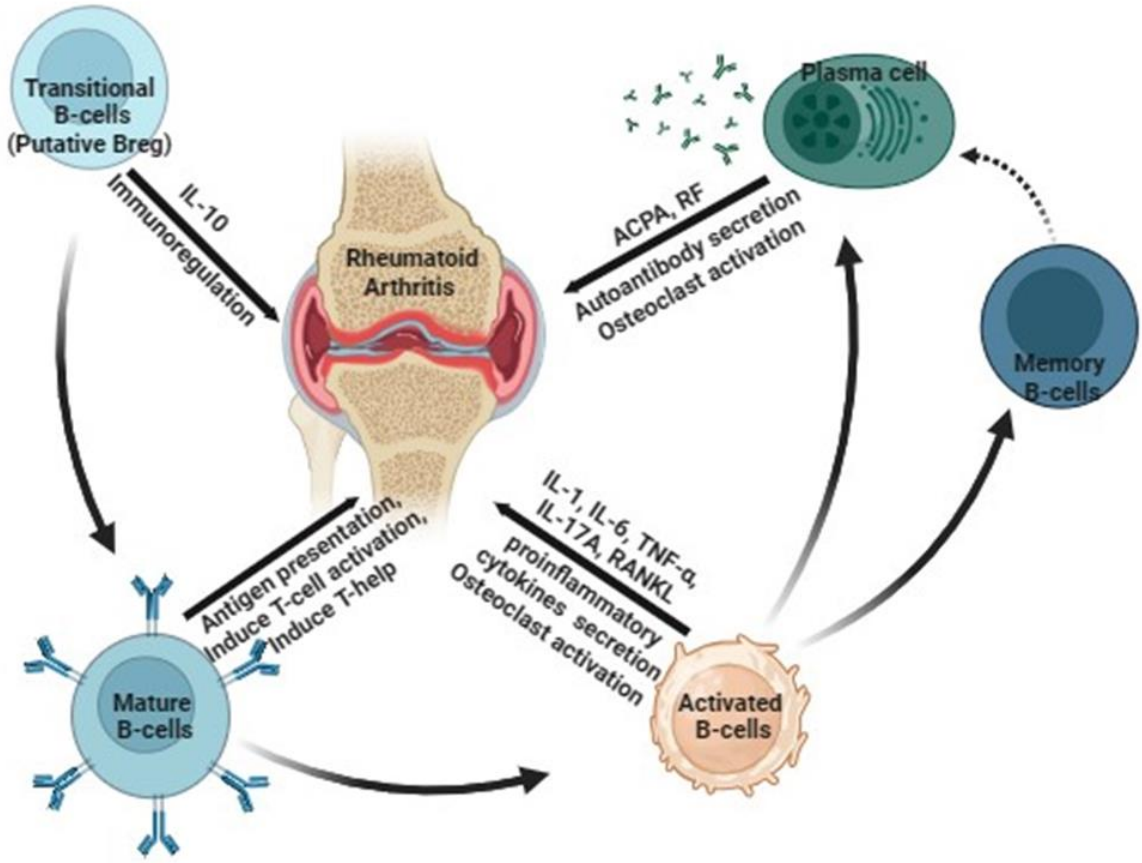


Figure 5 : B-cell subsets involvement in RA .

Briefly, in RA;

- Interactions between B- and T-cells lead to the activation and maturation of plasma cells, which produce autoantibodies[261].
- Activated B-cells assist T cells and drive the differentiation of effector T cells that generate proinflammatory cytokines [261].
- Moreover, B-cells influence various immune and non-immune cell activities by releasing cytokines like interleukin (IL)-1, IL-6, tumour necrosis factor (TNF)-α, and IL-17A. The production of proinflammatory cytokines and receptor activator of nuclear factor κB ligand (RANKL) by activated B-cells, and other cells promotes the activation and differentiation of osteoclasts, leading to bone resorption[262-269]
- B-cells contribute to bone homeostasis by facilitating the differentiation of mononuclear cells into osteoclasts through autoantibodies targeting citrullinated vimentin[265, 269].
- B-cells exhibit immunoregulatory functions by supplying IL-10 and employing other yet undiscovered mechanisms[261, 270-272]

Macrophages

Macrophages are recognised as one of the most abundant cell types in RA synovium [273]. Most macrophage resident-cells in synovial tissue have an active phenotype and drive joint erosion and cartilage degradation via multiple pathological responses involving the secretion of pro-inflammatory cytokines (IL-1, IL-6, TNF-α) [148, 150, 274] and chemo-attractants (CCL2 and IL-8) enhancing leukocytes infiltration by enabling angiogenesis and upregulation of degrading enzymes (proteases) as well as promoting fibroblast activation and proliferation [275-277]. On the other hand, IL-10 and TGF-β, the two primary anti-inflammatory cytokines are proposed to stimulate macrophages to control tissue inflammation and support angiogenesis, tissue remodelling, and tissue repair. The imbalance between the 2 macrophage subsets has been implicated in the development of RA [278, 279] .

As an APC, macrophages have been shown to exhibit significant levels of HLA-DR and leukocyte adhesion molecules, enabling them to engage in T-cell activation alongside B-cells in RA [265].

Mechanistically, osteoclast differentiation from macrophages precursor (monocyte) is the result of response to receptor activator of nuclear factor κB ligand (RANKL) [280, 281].

As such macrophages/monocytes are believed to be the cells fuelling the cycle of proinflammatory signals maintaining the disease. These pathogenic features are the main target of treatment in RA resulting in macrophages returning to resting state [282] supporting their critical role in RA as major player in the inflammatory state.

Synoviocyte-like fibroblasts

Joint destruction in rheumatoid arthritis involves several key processes leading to cartilage degradation. These include the expansion of the synovial membrane through pannus formation, characterized by the proliferation of resident cells, called synoviocyte which are fibroblasts like cells [149, 283, 284]. These are responsible for an increased expression of extracellular matrix-degrading enzymes, which contribute to the degradation of the protein in the cartilage matrix. These mechanisms collectively mediate the gradual deterioration of the joint. Additionally, synovial fibroblasts play a significant role in driving the pathogenesis of RA, further emphasizing their importance in disease progression as they are a major source of pro-inflammatory cytokine expression (such as IL-6, TNF and more) as well as chemo-attractant (SDF1/CCL12 for example) driving the influx of immune cells into the joint [283-286].

Other Blood Cell Subsets

Recent research has demonstrated the importance of the innate immune system in the onset and progression of RA. Natural killer (NK) cells and dendritic cells (DCs) are innate immune cells that may be involved in the initiation of inflammation that occurs in RA patients [236, 287] . The adaptive immune system, which is crucial in the later stages of the disease, is also activated by these innate immune cells [277].

NK cells are a subset of lymphocytes that account for 5 to 15% of the circulating lymphocytes. They are crucial elements of the early innate immune response that are involved in a spontaneous cell-mediated selective cytotoxicity without prior sensitization (before antigen priming) [288-290]. The $CD56^{dull}$ (termed mature cytotoxic NK) and $CD56^{bright}$ (termed immature NK) subsets NK cells can be

distinguished [277, 291]. The $CD56^{bright}$ favours dendritic cell maturation, improves CD4+ T-cell functionality, and promotes naive T-cell differentiation into Th1 subset by IFN-gamma secretion. They accumulate in inflamed synovium and engage in interactions with macrophages or their precursor cell (monocytes), thereby increasing the inflammatory response [292, 293]. When stimulated with IL-12, IL-15, or IL-18, $CD56^{bright}$ can, in a contact-dependent manner, increase TNF production by CD14+ monocytes [291].

NK cells expressing Granzyme B also contribute to the development of autoimmunity and to cartilage damage. The highest expression of granzyme B in synovial tissue of patients with early RAs strengthens the involvement of NK cells in onset and progression towards RA development [294].

DCs can play a complex and contradictory role in RA synovium [236, 295, 296]. DC can become polarised towards inhibition or activation which results in inducing tolerance or autoimmunity, respectively depending on the different signals they receive from the tissue microenvironment [277]. The creation and maintenance of Treg cells as well as the induction of T-cell unresponsiveness (anergy) are two mechanisms by which DCs can promote tolerance in the presence of immunomodulatory molecules, such as glucocorticoids, prostaglandins, or retinoic acid, in the environment . These molecules can inhibit DC maturation, cytokine production, and co-stimulatory molecule expression, leading to the induction of tolerance [297, 298]. Alternatively, in the inflamed context, the majority of APCs present are fully differentiated DCs (notably, conventional DCs and plasmacytoid DCs) which express class I and class II MHC molecules as well as T-cell co-stimulatory molecules, they may aid in educating and promoting the differentiation of self-reactive T-cells into effector cells due to

their antigen-presenting capability thereby contribution to the development of autoimmunity as seen in RA [236, 277].

Evidence supports the concept that fibroblasts cells [275, 286], are part of a complex cellular network with T and B-cells mediating RA pathogenesis.

Not much has been reported as to the frequency change of DC and NK in the blood of RA patients as well as with respect to clinical outcome. This highlights the need to explore in more detail the specific roles of these cells in RA to establish their potential as a predictive biomarker with respect to disease progression across IAC.

1.4 Inflammatory Arthritis Continuum (IAC) and disease management

The understanding of inflammatory arthritis (IA) has evolved substantially, and RA is now thought of as a continuum, involving a series of gene-environmental interplay leading to disease progression via multiple interacting pathways [299, 300]. The role of established risk factors in RA pathogenesis is fully described in **section 1.2.3.** Briefly, evidence of selected gene–environment interactions in RA development have been proposed, notably for smoking and HLA-SE loci [80, 96, 97] which are however related to T-cell immune mechanisms and antigen presentation [8, 88].

Finally, autoantibodies have long been associated with RA, notably RF since the ’60s. However, The discovery of ACPA revolutionised the IAC concept about 15 years ago. ACPAs are a heterogenous group of autoantibodies with capacities to recognise post-translationally modified citrullinated proteins [301]. ACPA-seropositivity confers an increased risk of developing RA in healthy individuals and therefore, is now recognised as a major European Alliance of Associations for Rheumatology (EULAR) classification criterion for RA [302].

The autoantibody shows high specificity (98%) and is found to be present in about 50% of established sever RA [303, 304], but relatively absent in non-RA (up to 6% in some diseases) [305-307].

The IAC is now a well-accepted disease progression with distinct stages towards RA [5, 308, 309] **(Figure 6)**. There is a long (up to 15 years) pre-clinical phase (also termed at-risk phase) involving a few stages (i – genetic risk to ii-environmental risk), leading to a break in tolerance and the development of systemic autoimmunity (iii) [302] manifested by the presence of autoantibodies (particularly, ACPA) [303, 310, 311] but no evidence of joint involvement [302].

Then comes (iv), a phase of arthralgia (joint pain), still, with no evidence of inflammation. This period can be followed by progression to clinical synovitis (v) (persistent or not), which over a few more weeks/months progress further to meet the criteria for RA diagnosis (vi) in a proportion of individuals[312] or to other types of IA [28] if the patients are referred to this stage of the IAC via an early arthritis clinic (EAC). At this stage, first-line therapeutic intervention can be initiated with conventional synthetic disease-modifying anti-rheumatic drugs (cs-DMARDs e.g. Methotrexate (MTX) in the first instance) with the aim of either remission which can be maintained for a long time with moderate drug adjustment and acceptable quality of life. However, in ~50% of cases, progression continues to a new stage (non-remission, vii) associated with resistance to cs-DMARDs and the need to resort to biological therapies (b-DMARDs).

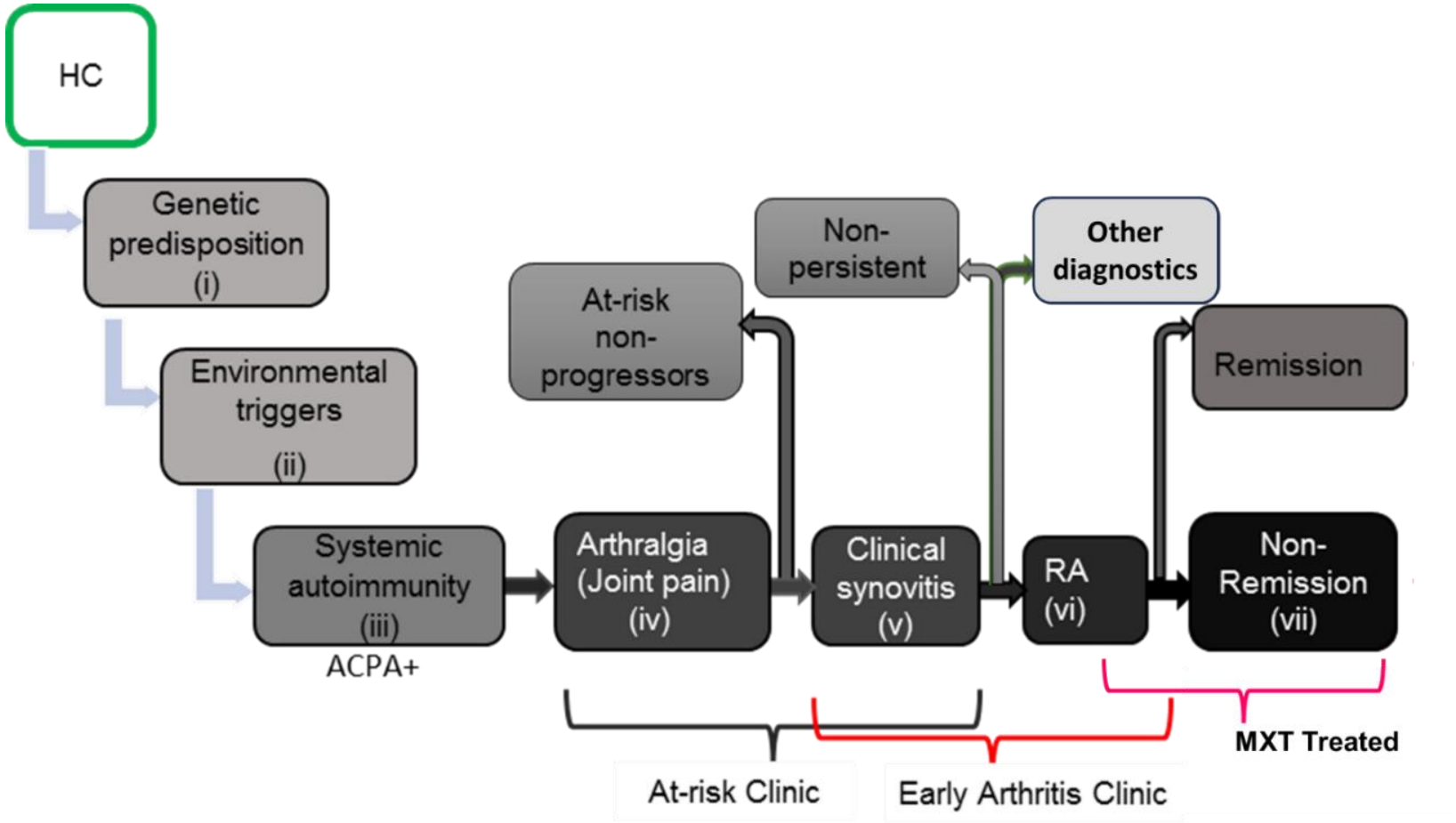


Figure 6 Conceptual Framework of Inflammatory Arthritis Continuum (IAC).

Adapted from [5, 308, 309, 313]. Other diagnostic included Osteoarthritis (OA); Psoriatic arthritis(PsA); Spondylarthritis(SpA)

1.4.1 Clinical management of RA

The presence of ACPA is key in identifying at-risk individuals although, only ~40% of ACPA+ at-risk individuals progress to RA [32, 312, 314] while ACPA-positivity heavily contributes to the diagnosis of RA [302] only approximately 50% of patients are seropositive at the first visit to a Rheumatology Clinic, leaving an unmet need for diagnostic biomarkers for ACPA-negative patients.

After diagnosis, the early control of inflammation is key to achieving remission and biomarker of response to first-line treatment (i.e. following National Institute for Health and Care Excellence (NICE) guidelines [32] would be essential to stratify patients for a classic approach with synthetic disease-modifying antirheumatic drugs (s-DMARDs) if they could be predicted to achieve remission or a more aggressive therapy with biologics (b-DMARDs) otherwise.

Following the achievement of drug-induced remission, the stability of remission can be quite variable, and predicting flare would be advantageous. The cost of therapy as well as the side effects of risk on patients drove the development of tapering treatment guidelines that also would benefit from biomarkers for successful tapering.

The management of RA currently relies on composite scores constructed from clinical observations, notably the disease activity score (DAS)-28, with 3 objective components (2 counts of tender and swollen joints, blood inflammatory markers which include C-reactive protein (CRP) or erythrocyte sedimentation rate (ESR)) and 1 subjective score provided by the patients (global health, on a scale of 1-100) [32, 312, 314].

DAS is used to monitor the magnitude of joint symptoms and level of inflammation in RA and different DAS scores and the formulae for calculating these scores have been described [314, 315]. In RA management, DAS28-CRP is widely used in making clinical decisions. EULAR response criteria categorised patients into 3 categories based on individual patient's responses to therapy measured by DAS28 score-

- DAS28 score >5.1 is indicative of high disease activity (HDA) and is used to determine eligibility for bDMARDs in the UK
- DAS28 score ≥ 3.2 but < 5.2 is considered moderate disease activity (MDA)
- DAS28 score < 3.2 is suggestive of low disease activity (LDA)
- DAS28 score < 2.6 defines a state of clinical remission [315, 316].

The DAS28 greater than 3.2 is commonly used as a threshold for classifying active RA patients and a target for tight control treatment [62, 316-319].

Of note, DAS28-ESR score of 5.1 has been shown to correspond to 4.6 for DAS28 score incorporating C-reactive protein (CRP), defining >4.6 as a threshold for HDA when using DAS28-CRP [320]. This suggests a need for strict adherence to the use of a specific type of measure in clinical decision-making because interchanging both scores may likely underestimate the number of patients with HDA as well as exclude those who might be eligible for intensive treatment with biologics if eligibility is based on DAS28-CRP score.

The DAS28 score was initially developed as a measure the severity of the symptoms and has since been widely employed in research and clinical settings [321-325] . However, its utility in assessing disease activity in RA patients remains challenging due to inherent confounders in each parameter of the DAS28-CRP [323, 326]. For example, a patient can be considered in remission

while still having 1 swollen joint or a raised CRP. Despite its endorsement by the European Medicines Agency as the preferred tool for determining remission in clinical trials, evidence suggests continued radiographic progression in patients achieving this goal [327, 328] . This raises questions about the suitability of the DAS28-based outcome measure utilized in research/clinical settings with respect of remission while its utility for assessing active disease was well accepted.

The subjective components of the original DAS28, (Ritchie articular index and 44 swollen joint count, along with patient-reported outcomes using general health visual analogue scale), reflect on its capacity to accurately discriminate between periods of high and low disease activity [329] , particularly as the subjective components tend to carry more weight in the score calculation and are susceptible to psychological variability [329] .

To address these limitations, alternative measures of disease activity , particularly those reflecting synovitis levels more accurately such as a two-component DAS28 (a score based on SJC28 and CRP alone), may offer improved performance in assessing treatment response compared to the original DAS28 or its components [330]. For instance, the 2C-DAS28CRP, has shown promise in studies with a research goal related to the pathophysiological manifestations of RA, demonstrating a stronger association with radiographic damage than the conventional DAS28 [325, 330] .

Replacing the conventional DAS28 in clinical trials or clinical daily practice should be approached with caution, considering the need to evaluate all core set areas as per the “Outcome Measures in Rheumatology” framework [331] .

1.4.2 Remission as the new goal of therapies

Due to time dependant increase in burden, the need to control RA disease activity early was identified and is now part of an overall modern strategy for managing RA, and particularly the goal of inducing remission as early as possible which has become the new target of therapy.

The first out of the ten recommendations on the treat-to-target (T2T) approach based on evidence and expert opinion set the primary target for RA treatment to be achieving a state of clinical remission, defined as the absence of signs and symptoms of significant inflammatory disease activity [332]. Where this seems impossible, a low disease activity state has become an acceptable alternative therapeutic goal, particularly in established long-standing disease [26, 313]

Although there is no cure for RA, early treatment using T2T approach has proven to reduce the risk of accumulating joint damage and to limit the long-term impact of the disease. The principle for treating RA early stemmed from the recognition of the concept of a “window of opportunity” [333, 334] which involves the immediate initiation of treatment upon diagnosis with a minimum delay to achieve early suppression of inflammation, before irreversible joint damage occurs in RA [313]. Documented evidence has validated this principle by demonstrating that early intervention in RA led to optimal therapeutic outcome and a lower disease burden [335-339]. For example, data from the Dutch Rheumatoid Arthritis Monitoring (DREAM) study [340] showed that patients at 1st referral with early RA (i.e less than 24 months duration of symptoms, drug naive) in a T2T protocol achieved remission faster with a lowered disease activity compared to those treated without such a target (median time of 25 weeks versus 52 weeks). Other studies have confirmed the evidence-based merit of T2T over conventional

treatment in terms of increased rate of remission [341-343], and/or achieving low disease activity [344-347] as well as being more cost-effective than usual care [348].

However, to initiate early treatment intervention with the T2T disease-modifying therapy, and prevent disability and deformity, there is the need for early diagnosis which sometimes, is missed due to classification errors given the heterogeneity in the presentation of RA [313]. Thanks to a collaborative task force between the American College of Rheumatology (ACR) and EULAR, a revised set of classification criteria was put forward in 2010 **(Table 2),** emphasizing RA characteristics that emerged early in the disease course, to allow diagnosis at the earliest possible stage [302].

Combining the new criteria with prompt referral was suggested as the key to achieving remission [313], the new treatment goal to prevent irreversible joint damage of RA.

Table 2 Revised ACR/EULAR rheumatoid arthritis classification criteria (Score-based algorithm)

A. Joint Involvement	score
1 large joint	**0**
2-10 large joint	**1**
1-3 small joints (+ or – large joints)	**2**
4-10 small joints (+ or – large joints)	**3**
>10 joints(at least 1 small joint)	**5**
B. Serology (at least one test result needed for classification)	
RF- and ACPA-	**0**
Low RF+ or low ACPA+	**2**
high RF+ or high ACPA+	**3**
C. Acute-phase reactant(at least one test result needed for classification)	
Normal CRP and normal ESR	**0**
Abnormal CRP or abnormal ESR	**1**
D. Duration of symptoms	
< 6 weeks	**0**
>6 weeks	**1**
Classification of patients as having definite rheumatoid arthritis	**≥6**

Anti-citrullinated protein antibody (ACPA), C-reactive protein (CRP), Erythrocyte sedimentation rate (ESR), and Rheumatoid factor (RF). Of note, the target population (individuals that should be tested) are those who have at least one joint with definite clinical synovitis(swelling) not better explained by another disease. Adapted from [302].

1.5 Clinical need in RA necessitating novel biomarkers

1.5.1 At-risk progression

Given the success of treatment for early RA [349], predicting an individual's progression to RA may enable preventive interventions. ACPA and/or rheumatoid factor (RF) are widely used to identify such individuals, and various prediction models were established by combining demographic, genetic, clinical, and imaging data [312, 350]. However, the rate of progression to IA is approximately 30-40% depending on the criteria used to identify at-risk individuals [351, 352]. Biomarker such as the presence of ACPA is currently the only key in identifying at-risk individual although, only approximately 40% of ACPA+ individual progress to RA [312]. In ACPA+ at-risk individuals, three CD4+T-cell subsets (naïve, Treg, and IRC) from our group showed good predictive value as biomarkers for progression to IA, both individually and when combined with clinical variables [3, 5].

1.5.2 Diagnostic

The 1987 revised classification criteria proposed by the American College of Rheumatology [353] are widely used for RA diagnosis **(Table 3),** allowing established RA to be differentiated easily from other forms of inflammatory arthritis diseases. However, it should be noted that the ACR was not designed to detect RA in its early stage being developed in populations with chronic/long-standing disease [354]. The 2010 criteria proposed by a joint ACR/EULAR committee, therefore accounted for this gap with improved overall sensitivity over the 1987 criteria [355]. Even though ACR/EULAR criteria determine synovitis based only on clinical examination, insisting on the presence of at least one joint with definitive synovitis has substantially led to classification in the earlier stage

of RA diseases thereby permitting an opportunity for early intervention using DMARDs.

The most typical clinical presentation of RA, in about 75% of patients, is the insidious onset of symmetrical polyarthralgia involving the small joint of the hands and feet characterized by tender joints as well as swollen joints [356].

ACPA-positivity contributes to the diagnosis of RA [302] although with only 50% seropositivity, leaving an unmet need for diagnostic biomarkers for ACPA-negative patients.

Table 3 The 1987 revised criteria for the classification of rheumatoid arthritis[353]

Criterion	Definition
1. Morning stiffness	Morning stiffness in and around the joints, lasting at least 1 hour before maximal improvement
2. Arthritis of 3 or more joint areas	At least 3 joint areas simultaneously have had soft tissue swelling or fluid (not bony overgrowth alone) observed by a physician. The 14 possible areas are right or left PIP, MCP, wrist, elbow, knee, ankle, and MTP joints
3. Arthritis of hand joints	At least 1 area swollen (as defined above) in a wrist, MCP, or PIP joint
4. Symmetric arthritis	Simultaneous involvement of the same joint areas (as defined in 2) on both sides of the body (bilateral involvement of PIPs, MCPs, or MTPs is acceptable without absolute symmetry)
5. Rheumatoid nodules	Subcutaneous nodules, over bony prominences, or extensor surfaces, or in juxta-articular regions, observed by a physician
6. Serum rheumatoid factor	Demonstration of abnormal amounts of serum rheumatoid factor by any method for which the result has been positive in 4% of normal control subjects
7. Radiographic changes	Radiographic changes typical of rheumatoid arthritis on posteroanterior hand and wrist radiographs, which must include erosions or unequivocal bony decalcification localized in or most marked adjacent to the involved joints (osteoarthritis changes alone do not qualify)
For classification purposes, ➢ a patient shall be said to have rheumatoid arthritis if he/she has satisfied at least 4 of these 7 criteria. ➢ Criteria 1 through 4 must have been present for at least 6 weeks. ➢ Patients with 2 clinical diagnoses are not excluded. Designation as classic, definite, or ➢ probable rheumatoid arthritis is not to be made.	

Adapted from 1987 criteria(ref). PIP= Proximal Interphalangeal MCP= Metacarpophalangeal, MTP= Metatarsophalangeal

1.5.3 Treatment

The goal of the treat-to-target strategy in RA is to achieve clinical remission (DAS28≤2.6) or at least a reduction in disease activity (DAS28<3.2) according to NICE guidelines thereby ameliorating the systemic and inflammatory debilitating effect of the disease [62, 317, 318]. RA treatment outcome has been greatly improved because of the major focus on early disease treatment.

The introduction of effective, although extremely costly, b-DMARDs, notably anti-TNF agents and more recently IL-6 blockade, co-stimulation blockade or JAK inhibitors requiring a DAS >5.1 for eligibility in the UK helped greatly in patients with long-lasting RA in which cs- DMARDs are not effective. But the introduction of biologics therapies in the early disease stage is currently not endorsed by NICE as it could impact long-term outcomes in RA. However, treating with a biologic without consideration of patients who would have responded to the 1st line therapy, with MTX would be extremely costly [357] and selecting which patient would benefit most and for which treatment option is still a daunting challenge.

The availability of different treatment options with different modes of action **(Figure 7)** and various response rates **(Table 4)** has also lent credence to the heterogeneity of the disease being a major factor to take into consideration. The “one for all” solution not fitting in RA is now being increasingly acknowledged.

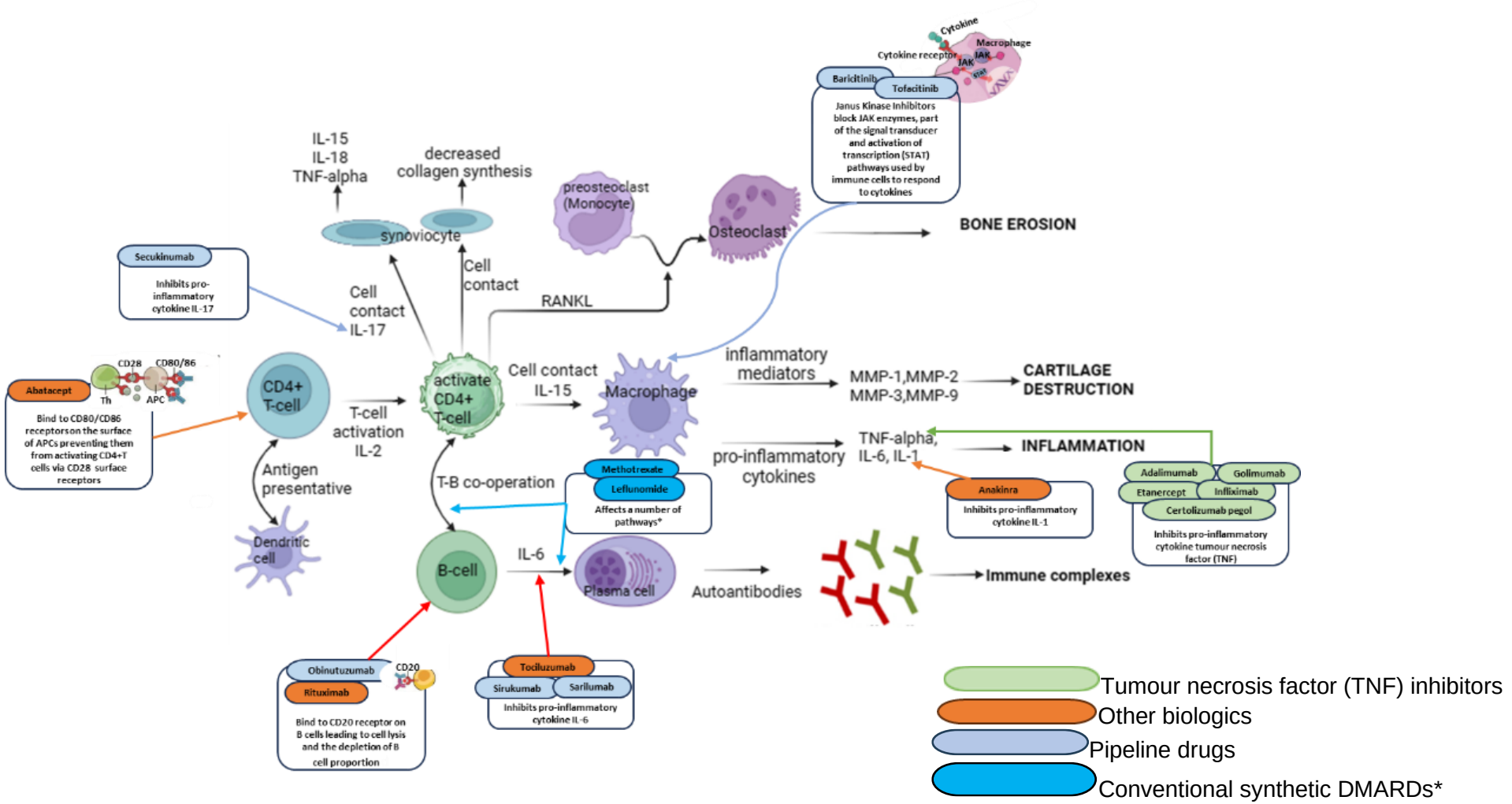


Figure 7 Mechanism of action of biologics DMARDs

*Mechanism of action of Methotrexate in RA has been extensively reviewed in [358] to include effects on several pathways of cellular metabolic functions: antagonism of folate-dependent processes, stimulation of adenosine signalling, inhibition of methyl-donor production, generation of reactive oxygen species, downregulation of adhesion-molecule expression, modification of cytokine profiles and downregulation of eicosanoids and matrix metalloproteinases (MMPs).Figure created with BioRender.com

Table 4 Response rate of common therapies used in RA treatment

therapy	Response rate at 6months	ref
MTX monotherapy	27.3% to 56.4%	[359-363]
MTX monotherapy	28%	[364]
RXT	21% to 54.3%	[365-369]
Anti-TNF	≤42%	[370]
anti-TNF + MTX	<50%	[371]

MTX= Methotrexate, RTX=Rituximab, anti-TNF= anti-tumour necrosis factor

Such differences between drug responses may underscore differences in underlying immunological pathways of inflammation, genetics, and other yet unknown mechanisms that might determine different paths to therapeutic responses in RA patients.

The variation in clinical response has also highlighted a gap in the therapeutic strategies to manage RA particularly how to identify patients who may likely respond to cs-DMARDs notably, MTX 1st line treatment without primary exposure to the adverse effect of biologics while allowing major economy.

Although, our group has demonstrated the predictive value of cellular biomarkers (notably multiple CD4+T subsets) for progression across the IAC, profiling of other immune cell types at all stages of the disease is yet to be fully explored and this thesis seeks to bridge this gap by investigating the value of a comprehensive Flow cytometry analysis of multiple LS simultaneously, for possible additional value as novel biomarkers for prediction of outcomes across the IAC.

1.6 Biomarker studies of RA prognosis

Biomarker studies encompass a broad spectrum of approaches, including genetics, transcriptomics, proteomics/lipidomics/metabolomics and more recently epigenetics, while immune cell phenotyping remains more constrained to immune cell diseases or residual cancer cell burden Biomarkers for RA, like in many other diseases have been extensively searched.

1.6.1 Genomics biomarkers across the IAC

Genome-wide association studies (GWAS) have unveiled a plethora of genes and genetic loci linked to both susceptibility and severity of rheumatoid arthritis (RA), offering crucial insights into the genetic underpinnings of the disease [372, 373]. Among them, the human leukocyte antigen (HLA) system stands out as the main genetic biomarker for RA. Particularly, *HLA-DRB1-SE* alleles have shown robust associations with RA susceptibility and severity, indicating increased risk and more severe disease manifestations [153]. Subsequent investigations have delved into the intricate interplay between specific *HLA-DRB1* alleles, autoantibody production, disease severity, and treatment response [374-377].

Polymorphisms within the *TNF* and *IL-6* genes have been scrutinized beyond their impact on RA susceptibility and severity, and genetic variations in the promoter regions have showed associations with response to anti-TNF therapies [378-380]. Other numerous genetic biomarkers showed association with poor treatment outcomes in RA, however their clinical utility remains modest as no single marker alone is predictive enough to guide treatment decisions [86, 88, 381-384].

Recent efforts combining data from large cohorts have identified genetic variants associated with response to MTX, a cornerstone therapy in RA management [381, 384]. These findings underscore the potential of genetic biomarkers in

tailoring personalized treatment strategies, however, it's crucial to acknowledge that they represent only a fraction of the molecular landscape shaping RA and as such have limited biomarker value.

1.6.2 Transcriptomic biomarkers across IAC

Transcriptomics, which analyzes RNA transcripts, provides crucial insights into the molecular mechanisms driving diseases like rheumatoid arthritis (RA). Dysregulated gene expression in RA contributes to inflammation, joint damage, and other pathological processes. Identification of transcriptomic biomarkers showed promise for early diagnosis (although not an area of clinical needs as mainly associated with ACPA+ disease), prognosis and treatment monitoring while not predicting of response to treatment so far. Notably, a study employed transcriptomic profiling to identify biomarker candidates in RA [385]. They examined gene expression patterns in synovial tissue samples from RA patients and healthy controls, pinpointing differentially expressed genes associated with disease activity and progression, particularly those involved in immune response regulation, cytokine signaling, and tissue remodeling. However, the applicability of using synovial tissue limit the value of this data as biomarker.

Afroz et al., conducted a meta-analysis integrating transcriptomic data from multiple RA studies to uncover robust biomarker signatures linked to disease severity and treatment response [386]. They established a potential transcriptomic biomarker (Signal transducer and activator of transcription 1 and 3 (STAT1/3) and Interferon regulatory factor 7 (IRF7)) for predicting clinical outcomes and guiding personalized treatment strategies for RA patients.

Several other studies have identified transcriptomic biomarker candidates [387-390]. Recently, a study utilised transcriptome-wide gene expression profiling to

identify early molecular signatures indicative of successful treatment outcomes, providing insights into the molecular mechanisms underlying TNF-inhibitor therapy efficacy, which may aid in the development of novel therapeutic strategies [389]. This finding is yet to be validated and replicated in larger and more diverse patient cohorts.

Integration of clinical parameters, other biomarkers, and imaging studies is still needed to demonstrate clinical utility. A panel of differentially expressed genes in RA synovial tissue, shedding light on immune cell activation and cytokine signaling pathways has been identified [391, 392] again with limited utility in daily practice. Recent advancements in single-cell RNA sequencing have provided deeper insights into cellular heterogeneity and gene expression dynamics within inflamed joints of RA patients [393, 394]. The findings collectively offer an overview of single-cell transcriptomes within CD4+ T-cell subpopulations implicated in autoimmune diseases. Integration of datasets from single-cell analysis still remain the biggest challenges for clinical utility.

Transcriptomic biomarker candidates often necessitate validation in independent cohorts while functional studies to confirm their relevance are important from a pathogenesis point of view. Altogether, validation and exploitation of these data is lacking and represent a major challenge because they use different recruitment for population, measure of the response, biological material, and more, rendering all data interesting but not useful so far in clinical practice. Integration with other omics data has been attempted for translating transcriptomic findings into clinical practice effectively but add another layer of complexity that does not help towards making this into robust biomarkers. Despite these challenges, transcriptomics still probably holds significant promise for identifying biomarkers and advancing precision medicine in inflammatory arthritis.

1.6.3 Proteomic Biomarkers across the IAC

Proteomics, which offers insights closer to the molecular mechanisms [395], focuses on quantifying mainly circulating (soluble) proteins to elucidate biological processes and decipher protein-dependent mechanisms governing regulatory mechanisms within living cells [396, 397]. Proteins serve as pivotal signals in cellular functions, and alterations in their expressions or activities can profoundly impact various cellular processes [398, 399]. Advanced technologies like liquid chromatography-mass spectrometry (LC-MS) have identified biomarkers with heightened sensitivity and specificity, offering promise in disease management [400]. Proteomics emerges as a promising avenue for identifying circulating biomarkers crucial for diagnosing, prognosticating, and monitoring treatment responses across this spectrum [401-404].

Proteomic analysis, across tissues, provided insights into disease-related protein alterations, stress responses, and therapeutic effects, offering information unattainable through other methods [400, 404-406].

Numerous proteomic studies on RA samples, including peripheral blood cells, synovial tissue, fibroblasts, and serum, have unraveled insights into RA pathogenesis [400, 404, 407, 408]. Proteomic studies have unveiled biomarker candidates in RA including cytokines (e.g., TNF-alpha, IL-6), acute-phase proteins (e.g., CRP, serum amyloid A), and autoantibodies (e.g., rheumatoid factor, anti-citrullinated protein antibodies) [408-412]. A protein signature with diagnostic potential to discriminate patients with PsA from RA have also been identified via this approach [400]. The most successful story for a clinically useful biomarker in RA is indeed that of ACPA [413, 414]. These were identified over 30 years ago as IgM autoantibodies present in many people, reacting to unknown antigens in skin tissues, while highly specific to RA when turned into an IgG [310,

415]. Their value as biomarkers took over 15 years to be recognized but they are now routinely assessed as a classification criteria for RA diagnosis. They further define a subset of RA cases with a more rapid and severe disease course [310, 416].

Proteomic signatures for proteins linked to innate immune mechanisms, like TNF, Toll-like receptor 2 (TLR-2), interleukin-1A (IL-1A), and interferon-gamma (IFN-G) have accurately predicted RA development in at-risk individual, suggesting innate immune involvement in the preclinical phase [411]. Alterations in the serum proteome preceding clinical RA onset underscore its potential as a biomarker tool for classifying at-risk individuals and unraveling molecular pathways implicated in RA development [411]. In non-RA forms of inflammatory arthritis such as PsA, biomarkers such as IL-17, IL-23, and matrix metalloproteinases (MMPs) reflect characteristic inflammatory and tissue remodeling processes [417-420]. Similarly, AS-specific biomarkers identified through proteomics (using tandem mass tag (TMT)-based quantitative proteomics) include the combination of molecules modulating inflammation (C-reactive protein (CRP)) and serum amyloid A1 (SAA1) as the best panel for the diagnosis of active AS [421]. As such that can aid in early differential diagnosis (and disease progression monitoring) allowing to separate RA from these 2 very similar conditions at the early presentation stage, notably as those are seronegative diseases and RA remains in need of biomarker for the half of patients that are ACPA [403, 407].

In conclusion, proteomics bridges the gap between genetic and transcriptomic information by revealing the functional state of proteins that also have possible methods of quantification for clinical use (ELISA for example). While proteomics has identified several RA-related candidates, rigorous validation and further

analysis of differentially expressed proteins are crucial steps toward translating basic scientific findings into clinical applications.

Lipidomics and metabolomics [422, 423] are still emerging technologies that have not yielded notable data in RA to date. Moreover, these technology often overlook posttranslational modifications of proteins while they may require to be taken in consideration in RA, some of which are crucial for their biological activities.

1.6.4 Epigenetic biomarkers across the IAC

Epigenetics has recently been brought at the forefront of the understanding of many disease pathologies [424-427]. In RA this mechanism has been studied and implicated in the development of the diseases [428-432]. The role of epigenetic changes is therefore well established in RA warranting the potential of epigenetic biomarkers as tools for RA management.

RA's complex etiology, reliant on the interplay of genetic and environmental factors, is closely intertwined with epigenetic changes linked to transcriptional regulation, reflecting pathogenic changes associated with disease states. Methylation studies have unveiled epigenetic modifications, influencing gene expression and disease phenotype, disease activity, severity, and treatment response in RA [431, 433-436] and spondyloarthritis (SpA) patients [437, 438]. Altered DNA methylation patterns have been observed in genes related to immune response and inflammation. For example, hypomethylation of the promoter region of the *CTLA4* gene [93] and other genes has been associated with RA susceptibility, increased disease severity, and progression [432, 439]. Recently, a panel of DNA methylation changes has accurately distinguished RA patients from healthy controls while more importantly being correlated with disease activity. Assays able to measure accurately and robustly DNA methylation have been developed, notably methylation specific qPCR (qMSP)

assays [427, 440, 441]. Based on data showing differential methylation in early RA [428], TRIMID group developed 2 qMSP assays for genes delivered in this pathogenesis study (IL-17 and TNF) and notably demonstrated their value as diagnostic biomarkers in RA (TNF and IL-17) [359, 428] and another study for response to MTX [442].

Other histone modifications, crucial for regulating gene expression by modulating chromatin structure, have also been implicated in RA [443, 444]. Increased histone deacetylase (HDACs) activity and altered histone methylation patterns have been linked to key inflammatory pathways in RA, influencing disease progression [443-445].

Enzymes involved in epigenetic modifications, such as DNA methyltransferases (DNMTs), histone acetyltransferases (HATs), and HDACs, are also potential biomarker candidates and therapeutic targets in inflammatory arthritis [443, 446-449], due to their dysregulated expression and activity contributing to aberrant epigenetic regulation and disease pathology.

In summary, epigenetic biomarker candidates in the inflammatory arthritis continuum offer valuable insights into disease pathogenesis and progression, as well as potential targets for precision medicine approaches. Integrating data from multiple biomarker types can enhance predictive models and facilitate personalized medicine approaches in RA and other inflammatory diseases.

1.6.5 Other Regulators of Transcription

Small or long non-coding RNA are a new family of regulators which have been shown to have important mRNA regulatory activity [450-455]. Dysregulated expression of specific microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) has been associated with disease activity, joint damage, and therapeutic response in RA and SpA patients [452] . Specific miRNAs, like miR-

146a [456-458] and miR-155 [459, 460], have been identified as potential biomarkers in RA due to their regulatory roles in inflammatory pathways. Dysregulated expression of lncRNAs has also been reported to contribute to RA pathogenesis by modulating inflammatory responses and immune cell function [461, 462] .

Furthermore, differential DNA methylation at specific gene loci has been correlated with response to TNF inhibitors (TNFi) and MTX in RA patients [439, 452], highlighting the potential of epigenetic biomarkers in predicting treatment response.

1.7 Review of existing blood cell subset biomarkers across the IAC

It is well known that additional types of lymphocytes including T-cells, B-cells, NK-cells, and their subsets are important players in the development of inflammation in RA and the disease progression across all phases in the IAC [194, 370, 463] . Immunophenotyping investigations of circulating blood lymphocytes are believed to be able to provide useful information that may reflect various states of the disease or suggest possible functionalities such as change in migration capacity of immune cells to various sites. This has led to the concept that blood cell phenotyping could have biomarker value for RA progression from the at-risk stage, diagnosis, as well as for treatment response.

1.7.1 Blood cell biomarkers of progression from the at-risk stage to IA

Research demonstrating either functional or numerical dysregulation within immune cell subsets particularly in the lymphocyte subpopulations in at-risk arthralgia individuals are limited. The current literature aimed to understand the value of lymphocyte dysregulation while highlighting their association with disease progression to IA from either seropositive (ACPA-positive) arthralgia

patients or seronegative arthralgia patients. However, data reporting lymphocyte dysregulation at this pre-clinical phase are not consistent and most studies focused on seropositive arthralgia patients (as shown in **table 5)** with evidence from seronegative patients being scarce due to the difficulty in identifying them.

Table 5: Summary of candidate blood cells as biomarker for progression to IA

cohort	Patient population	subset	study design	sample type	statistical analysis	ref
Amsterdam Reade ASP cohort	Patients with arthralgia (n=113)	Lower CD8+T-cells Lower memory CD27+ B-cells	cross-sectional study, observational	fresh blood	Univariate association (MWU test and T-test)	[464]
Groningem Very Early IA cohort	SAP (n=26) versus HC (n=24)	higher "putative" Th17 cells	cross-sectional study	fresh blood	univariate associations (MWU test)	[465]
	SAP (n=30) versus HC (n=41)	decrease in total NK cell numbers/NK56dim in SP arthralgia,		fresh PBMC	Univariate: ANOVA	[466]
Leeds at-risk cohort	ACPA+ individuals with pain but no clinical synovitis (n=103)	Reduced Naïve/Treg & increased IRC	cross-sectional study, observational	fresh blood	Modelling: multivariate logistic regressions (AUROC/OR) and Cox models	[3]
	ACPA+ individuals with pain but no clinical synovitis (n=103)	higher IRC and loss of Naive/Treg	cross-sectional study, observational	fresh blood	Modelling: multivariate logistic regressions (AUROC/OR) and Cox models	[5]

Analysis of Variance (ANOVA), Area under(AU) the receiver operating characteristic(ROC), Odd ratio(OR), Mann-Whitney(MWU) test, inflammatory-related cell(IRC), Anticitrullinated peptide antibody(ACPA), Peripheral blood mononuclear cell (PBMC), health control (HC)

CD8+ T-cells

Evidence from the Amsterdam Reade cohort showed that lower circulatory CD8+T-cells frequencies were associated with progression, within 24 months to the development of IA [464]. The explanation for lower cytotoxic CD8+T-cells frequency prior to progression remains unclear. But the authors suggested that it might be a reflection of increased immigration of these cells to synovium. At-risk arthralgia patients may have a mild infiltration of CD8+ T-cells in their synovium [184]. In pre-clinical stage, other reports also suggested that CD8+ T-cells make up ~40% the total T-cells infiltrating the synovium [184, 467]. Alterations in homing molecules expressed by CD8 T cells also suggested enhanced migration to the synovium that may also be relevant to the disease development [468]. It is therefore possible that CD8+T-cells migrate to sub-clinically inflamed tissue while alternatively, they could migrate to lymphoid tissue notably through CXCR3 signaling, accumulating in the marginal zone of lymph-nodes [464], An argument was also proposed suggesting that the reduction of CD8+T-cells may be attributed to the elimination of autoreactive CD8+T-cells from the total CD8 pool via class I-restricted cross-presentation of self-antigens expressed in tissues outside the lymphoid compartment [469, 470].

Alternatively, following prolonged exposure to infectious agents, expression of homing molecules expressed on CD8 T-cells is altered [468] notably CXCR3 expression. This was used to suggest that reduction in CD8+T-cell cells was related to a population contraction resulting from cell-death following prolonged exposure to the infectious agent, proposing a role for subclinical infections as a novel environmental risk factor for RA. This was also aligned with suggestion in the past for a role for virus like Human parvovirus B19 [471], Epstein-Barr virus [472] and others [473].

Finally, cell-death causing the release of inflammatory mediators that also serve as destructive triggers (notably Netosis) were also proposed to explain the reduction in circulating CD8+T-cells [474].

Therefore, there remains a gap in knowledge as to whether CD8+ T-cells could have biomarker value based on all the possible functional role they may have in progression to RA.

B-cells

The Amsterdam Reade study also reported a lower frequency of CD27+ memory B-cells in patients who progressed to IA within ≤1 year notably compared to those who converted to IA [464] later. It was suggested that increased early migration of memory B-cell into the inflammatory site (synovium) may play a role in lowering circulating memory B-cells frequency. Evidence of elevated frequency of memory B-cells in synovium of early RA patients also underpins this hypothesis [475]. Alternately, memory B-cell might have homed towards the bone marrow or lymph node in line with a report of higher B-cells in the draining lymph nodes of inflamed joints in patients at the earliest phase of inflammatory arthritis [476]. However, the study did not investigate whether CD8, or memory B-cell subsets had any value for predicting progression to IA beyond their association with the outcome.

CD4+ T-cells

Beyond our own group's work, a study looking at regulatory T-cells (CD4+ CD25+ FoxP3+Tregs), reported no significant difference between seropositive arthralgia patients (SAP) who developed RA and non-progressors [477]. This was used to suggest that Tregs were not associated with progression to IA/RA. However, the phenotype used to define Treg at the time (CD25/Foxp3) was not complete (no

use of CD127 marker). CD127, also known as interleukin-7 receptor alpha (IL-7Rα), is an important marker in identifying and characterizing Tregs. In defining Tregs, CD127 is often used as a negative marker [478] . That means, among CD4+ T-cells, Tregs typically express lower levels of CD127 compared to conventional CD4+ T-cells. Therefore, CD127low or CD127− cells within the CD4+ T-cell population are often enriched for Tregs. The percentage of CD25+Foxp3+ Tregs expressing CD127 varies depending on the tissue, individual, and experimental conditions. Generally, a substantial proportion of CD25+Foxp3+ Tregs are CD127low or CD127−, indicating their regulatory phenotype [479]. In situations (like non-availability of FoxP3) where only surface staining can be performed, CD127 can indeed be used as a surrogate marker for Tregs in conjunction with other markers like CD25. However, it's important to note that while CD127 is associated with Treg phenotype, it is not a specific marker for Tregs. Foxp3 is the only accepted marker of Treg [480] . Therefore, using CD127 alone may not capture the entire Treg population accurately while its absence may slightly over estimate the frequency of Treg, particularly in a disease context where CD127 expression might be dynamically regulated [481] .

In contrast, our department demonstrated that in ACPA+ individuals without clinical synovitis, 3 CD4+T-cell subsets naïve CD4+T-cell, Treg and IRC showed good predictive value as biomarkers of the progression to IA, both as a score and in the regression model [3, 5]. The reason for this discrepancy is not clear but the Treg phenotype was not the same and no correction for age-associated change in Treg was applied to the SAP study.

To this date association and predictive values but not yet added value beyond current model of prediction for these 3 subsets were demonstrated statistically,

notably above using ONLY clinical [3, 312] and/or clinical + imaging data [312, 350].

NK cells

The biological role of NK-cells, notably from the CD56dim NK-cells subset, is a cytotoxic one associated with an anti-microbial role mediated by the secretion of IFN-γ. A reduction in the number of peripheral total NK-cells (and in the CD56dim subset) was observed in SAP, however, compared to HC [466]. Whether the lower frequencies of NK cell/CD56 dim subset preceded or is associated with the IA/RA development was not described as the study did not provide detail on whether the at-risk individuals progressed. Given that the comparison group was healthy individuals, it is possible that alteration in NK-cells/CD56dim subset may play a role in pre-clinical phase, however, this needs to be investigated.

Th17 cells

Data on Th17 cells from patients within the at-risk arthralgia phase is scarce, however, indirect data on a Th17 surrogate phenotype (defined as CD4+CD161+ T-cells) seems to indicate a possible increase in Th17 cells was associated with SAP [466]. Findings suggested that in the early immune events leading to clinical synovitis (i.e. progression to IA detected by joint swelling or inflammation), an elevation in numbers or frequency of Th17 was observed and may therefore play a crucial part. However, the CD4+ CD161+ cells phenotype includes various cell subsets such as IL-17 and IFN-y expressing cells, representing mixed populations of Th17, Th17/Th1, and Th1 cells. This nonetheless is an important observation that need to be further investigated.

One of the early events associated with RA pathogenesis is a change in the cytokine environment that promotes Th17 differentiation [482]. The “Translational Research in Immune-Mediated Inflammatory Disease” (TRIMID) group also showed that Th17 cell frequencies are reduced in early RA due to migration to the site of inflammation via CXCR4 expression [359]. Furthermore, increased serum levels of IL-17 proceeding progression toward clinical synovitis has indeed been reported in pre-RA patients [177], suggesting that such an event is likely, and further research needs to confirm such hypothesis.

1.7.2 Blood cell subsets as diagnostic biomarkers of progression from IA/UA to RA

Here I will discuss multiple immune cell types/subsets reported to play a role in the progression to RA from the IA stage with respect to their possible value as biomarkers as shown in **Table 6.**

Table 6 summary of candidate blood cell as biomarker for RA diagnosis

cohort	study population	subset	recruitment method used	sample type	statistical analysis	ref
Amsterdam Reade ASP cohort	early drug-naive RA (n=89) versus HC(n=37)	lower CD3 T cells lower activated CD56+CD3+ NKT-cells lower memory (CD27+) B cells with CD80+	cross-sectional study, observational	fresh blood	Univariate association (MWU test and t-test)	[464]
Groningem Very Early IA cohort	newly diagnosed RA patients (n=35) versus HC (24)	lower “putative” Th17 cells	cross-sectional study		Univariate association (MWU test or t-test)	[465]
	SP RA (n = 45) SN RA (n = 12) versus HC (n=41)	decrease number/potency in total NK cell lower NK56dim in SP RA.		fresh PBMC	Univariate: ANOVA	[404]
Leeds at-risk cohort	early inflammatory arthritis DMARDs naïve-patients (n=179) versus HC (n=49)	lower th17 higher expression of CXCR4	longitudinal prospective study	fresh blood	multivariate modelling	[359]
	evolving IA patients (n=294)	loss of Naive/Treg	cross-sectional study, observational	fresh blood	multivariate modelling	[3]
Japan early-onset RA cohort	early-onset RA patients (n=11) versus disease control (OA, n=6)	elevated th17 derived th1 cells		fresh blood PBMC	Univariate association (MWU test)	[483]
French RA cohort	RA (n=92) versus disease + healthy control(n=25)	lower Tregs	cross-sectional study, observational	fresh blood PBMC	multivariate regression analysis	[484]

Analysis of Variance (ANOVA), Area under(AU) the receiver operating characteristic(ROC), Odd ratio(OR), Mann-Whitney(MWU) test, , seropositive/negative (SP/SN), Peripheral blood mononuclear cell (PBMC), health control (HC**),** Rheumatoid arthritis(RA)

CD4+T-cells

My group reported on the lower frequencies of naïve CD4+T-cells, higher IRC [7] and lower Treg [485] in early RA. The potential as a biomarker value to diagnose RA was also established [5]. However, full modelling of the added value beyond association of 3 cell subsets as diagnostic biomarkers remains to be demonstrated rigorously using regression models.

Th17 cells

Evidence suggests a role for Th17 cells in early-onset RA patients [483, 486], Alteration in terms of lower peripheral numbers of Th17 cells (defined by the surrogate phenotype CD4+CD161+ T-cells) in early RA and subsequent enrichment of these cells at the site of inflammation (joint) suggest a role for these cells in the early immune events leading to clinical synovitis [465]. However, the limitation of this phenotype remains similar to what was discussed at the at-risk stage. Consistent with the study, TRIMID group demonstrated that fewer circulatory proportions of Th17 cells (not using flow cytometry assay, but a quantitative methylation-specific PCR (qMSP) biomarker assay of the epigenetic modification of the *IL-17* gene itself as a surrogate for the quantification of Th17cells) in individuals with early IA symptoms (DMARDs naïve, <12 months) was independently associated with progression to RA compared to patients who did not develop RA [359]. The study further used a predictive model resulting in 73% accurate diagnosis adjusted for all confounders notably in an ACPA-negative subgroup where Th17 and swollen joint count were the only predictors. Interestingly, and consistent with this report [465]. Th17 cells using a different surrogate phenotype (CCR6+ cells) showed a higher frequency in ACPA-negative RA patients with a shorter disease duration [487]. However, a model

using Th17 (defined as CD4+ CCR6+ cells) cells could not show predictive value for progression to RA in ACPA+ patients [487]. The author suggested an underlying difference in pathologic pathways involved in both cellular and clinical risk factors associated with developing RA in seropositive at-risk individuals. Alternatively, in line with data in SAP, Th17 implication may be an event that occurs earlier in ACPA+ disease compared to ACPA-.

Another study [483] used more refined phenotypes of the putative CD161+ cells Th17: Th17 cells IL17+CD161+ or Th1 cells IFN-γ+CD161+ or Th1/17 cells IL17+IFN-γ+CD161+, and showed increased of Th1 over Th17 cells in the peripheral blood of the early-onset RA patients, followed by the subsequent reduction in Th17 cells but not Th1 cells suggesting a role for Th17 cells in the early phase of RA. However, it remains to be elucidated whether this alteration has a predictive value for discriminating progression towards RA or other forms of inflammatory arthritis notably, AS or PsA where these cells are predominant.

Regulatory T-cell (Treg)

There is an abundant literature about Treg in RA [3, 488-496] however, it is difficult to reconcile it all, due to the use of multiple phenotypes (from single to multiple markers), RA being one of the 1st diseases to be investigated for immune regulation back in the early 2000's. As consensus developed over time, studies are now using what is accepted as a definite phenotype for Treg (CD4+CD25^{high}Foxp3+CD127^{low}) [8].

Patients with RA have significantly less circulating Treg compared with healthy controls [272]. Treg frequency at baseline was reported to be lower in RA

compared to OA or other mechanical diseases (free of systemic inflammation) [484].

TRIMID research group was instrumental in defining the lower circulating Treg frequency in early RA [485] and later [3, 5] as well as establishing the age-relationship in health and its loss in early RA [4]. However, as reduced frequency of Treg predated the new-onset rheumatoid arthritis as demonstrated by my group [5], this suggests a pathological role for Treg dysregulation towards the development of RA, at least in ACPA+ disease. A definite value as diagnostic biomarker remains to be established (also comparing ACPA+ and – diseases), while insight was provided again by my group in a recent paper suggesting association with RA but NOT yet the predictive value of Treg for diagnosis [5].

B-cells

Regulatory B-cells (Breg) are a recently identified subset of B-cells, notably producing IL-10 [497-502]. Reduced frequency of regulatory B-cell was inversely associated with disease activity in patients with new-onset RA [272]. However, whether this down-regulation of circulating Breg is associated with predicting progression to RA is yet to be determined.

NK cells

In the study mentioned above in SAP [466], a decline in NK cells and NKCD56^{dim} subset was also reported in newly diagnosed, treatment-naive, ACPA-positive RA, suggesting that alteration in NK cells may contribute to RA development rather than represent the consequence of long-term inflammation. However, whether the alteration preceded RA was not clear from that study but the comparison with health control (HC) was used to present NK and NKCD56^{dim} cell

subsets as potential biomarkers for RA. There is little or no evidence showing the clinical value of NK cells/subsets alterations in relation to disease progression from IA to RA. This gap remains to be investigated.

Conclusion

Data investigating changes in percentages of various lymphocyte subpopulations between outcomes across the IAC are not robust yet due to lack of sufficient statistical power [483]. Furthermore, many studies could not be related to the context of progression to RA because they are not comparing the right population of patients in order to establish a diagnosis (i.e., they used healthy control and not patients with IA who do not progress to RA) [272, 464, 465, 484, 503-505]. Altogether, there remains a gap in phenotyping comprehensively lymphocytes and subsets implicated in RA pathogenesis in order to establish their possible value as predictive biomarker with respect to disease progression to RA.

1.7.3 Blood cell biomarkers of clinical responses to cs/b-DMARDs treatment in RA

Prediction of response to drugs, conventional synthetic disease-modifying anti-rheumatic drugs (csDMARD) is paramount to preventing long-term damages and disabilities notably, in early disease. It may use different cell subsets compared to biological drugs targeting specific aspects/cells of immune systems **(see Figure 7, page 44)**

- anti-TNFs (cytokine blockade against tumour necrosis factor)
- co-stimulation blockage (Cytotoxic T lymphocyte with associated antigen-4-Ig, CTLA4-Ig)
- B-cell depletion (anti-CD20 monoclonal antibodies), and
- IL-6 inhibition (anti-interleukin-6 receptor monoclonal antibodies)

These are likely to be predicted by different cell subsets, indeed due to targeting such cells themselves as well as these targets potentially having different roles at different stages of the disease. Biologics are readily available in the United Kingdom although very expensive with varying rates of efficacy in RA (reported between 40-65%) [365-371]. Therefore, these would also benefit from a more rational use based on selecting the right drug for the right patient at the right time being expensive with only about a 50% chance of success.

Here is discussed the multiple immune cell types/subset reported to show association with response to cs/b-DMARDs (as summarised in **Table 7 to Table 10),** with respect to their possible value as biomarkers.

1.7.3.1 Blood cell biomarkers of responses to Methotrexate (cs-DMARD) treatment

Methotrexate (MTX) remains the most widely recommended as the standard first-line csDMARDs therapy for patient with early RA [506-509] with best documented efficacy, safety, and relatively low costs [510-516]. csDMARDs therapy have been shown to effectively control systemic inflammation in patients with new-onset RA after 12 weeks of treatment [272], though this is still limited due to heterogeneous clinical efficacy that is furthermore, difficult to predict [364]. Immune cell subsets reported to show association with responses to MTX is shown in **Table 7.**

Evidence from a randomized blinded trial demonstrated that poor-prognosis (based on DAS28-ESR) of early RA patients on initial MTX monotherapy achieved good clinical response at 24 weeks without further need to step-up to combination therapy [364]. This suggests that a good proportion of early untreated patients (up to 30%) that would likely benefit from MTX monotherapy [364]. Nonetheless, poor clinical response occurs in many patients which take a prolonged period of time before MTX-nonresponse becomes evident [512-516] creating an opportunity for the disease to exacerbate in this group of patients. This highlights the need for a biomarker(s) to predict patients with best chance of benefiting from MTX and if otherwise, allow alternative therapy early enough to avert worse clinical outcome [517] in potential MTX non-responders.

Regulatory T-cell (Treg)

Treg (defined as CD4+CD25^{high}CD127^{low}FoxP3+) is believed to play an important role in RA. Initially, data compared various states of disease activity. For instance, a lower frequency of Treg (CD4+CD25+Foxp3+ T-cells) in new-onset RA patients was inversely associated with disease activity, although whether the reduction was predictive of response to treatment was not investigated [272]. In MXT-treated RA patients (after 6 months on therapy) increased frequency, and inhibitory activity of 2 phenotypes of Treg (intra-marker Treg defined as CD4+CD25+Foxp3+ and extra-maker Treg defined as CD4+CD25^{high}+CD127low−) were associated with clinical remission compared to active RA [518]. The Foxp3+ Treg phenotype showed a strong positive correlation with CD4+CD25high+CD127low− population, hence the authors suggest that either the intracellular (Foxp3) or extracellular markers only (CD25high/CD127low) can be used to define Treg. Whilst others have used a combination of both extra- and intra-cellular markers (CD4+CD25^{high}CD127^{low}FoxP3+) for Treg identification on MTX at BL [5, 519]. Data are, therefore, inconsistency across studies due to issue with variabilities in markers used to define Treg.

A low density of the CD39 cell surface marker on Treg (defined as CD39+CD4+CD25+FoxP3+) was only associated so far with their reduced suppressive activity as well as with unresponsiveness in RA patients treated with MTX [512]. Another study demonstrated a direct relationship between higher expression of CD39 on Tregs and their inhibitory activity and furthermore higher proportion of CD39+ Treg at baseline was shown to have a potential value as an independent predictor of good clinical response to MTX monotherapy after 4

months in a univariate analysis [363]. Whether CD39+ Treg remained a fair predictor of MTX responses (with fully adjusted clinical parameters) was not clearly demonstrated in the study. CD39 is an exonucleoside enzyme that is responsible for the production of adenosine (ADO), a crucial anti-inflammatory mediator of MTX action. A more recent report then progressed this work suggesting a biomarker value for CD39 expression on circulating Tregs in predicting MTX response [519], showing that higher levels of CD39 expression on Treg (MFI) predicted MTX-induced remission at 6 months (defined by DAS28CRP) with great specificity and sensitivity (AUC: 0.725), demonstrating a discriminative utility

In early untreated RA, phenotypic changes in Treg were studied in the course of methotrexate treatment (These were defined as

- FoxP3+CD25^{+},
- FoxP3+CD127-,
- CD25+CD1 percentages calculated based on analysis of CD4^{+} T cells)

and subpopulation of FoxP3 CD25^{+} :

- CTLA4+surface;
- CTLA4+intracellular,
- FoxP3+ICOS+;
- FoxP3+CD40L+;
- FoxP3+PD-L1[361]

The findings suggest that a higher % CTLA4 surface phenotype and FoxP3+CD25+Treg at baseline were independently associated (using MWU test)

with good clinical respond to MXT monotherapy after 24 weeks with 56% of patients achieving remission according to DAS28. This further contributed to the idea of value for Treg as a biomarker to predict MTX response.

Our own study using the accepted minimum Treg phenotype ($CD4+CD25^{high}FoxP3+CD127^{low}$) which allows better discrimination between recently activated T-cells (CD4+CD25+FoxP3+CD127+) and Treg, did not however, select Treg as a predictor (OR15.4 $p<0.0001$) of MTX-induced remission, although lower frequencies were weakly associated ($p<0.05$) with remission [5]. This work in 70 early RA patients demonstrates the need to use appropriate statistical approaches to establish the difference between an association with outcomes (using MWU test) for potential biomarkers as opposed to those that may achieve clinical utility with predictive values (using OR, AUC, and regression model).

Treg therefore remain a potential biomarker for DMARDs response, more likely when using additional phenotypic marker (such as CD39 or CTLA4) that could add value to multi-parameter model for predicting clinical response.

Naïve CD4+ T-cells

Our group was the first to describe the potential of naive CD4+T-cells to serve as a biomarker for clinical response to MTX treatment, evaluated using the EULAR response criteria, DAS28, distinguishing responder patients from non-responders [520]. A 2nd study [5] demonstrated the clinical utility of naïve CD4+T-cells across the IAC and replicated data showing the value of naïve CD4+T-cells frequencies at baseline for predicting MTX-induced remission (fully adjusted model for clinical parameters) as well as removing Treg and IRC from this model. In this first study however, higher naïve cell frequency was not found to be associated with

remission in patients who received MTX concomitant to anti-TNF, suggesting a unique role of the naïve subset in the prediction of MTX response in RA. As a result, employing naïve CD4+T cells as a biomarker may have positive implications for a personalisation, as well as cost-effectiveness of treatment in early RA patients starting MTX. As such LIRMM is currently undertaking a randomised clinical trial (Targeted Treatment Early With Etanercept + Methotrexate vs.T2T Care for DMARD-naïve Early RA Patients Based on naïve T-cell Stratification (TEEMS)) stratifying early RA patients

- predicted NOT to respond to MTX alone using naïve CD4+T-cells as biomarker, between 2 arms :
    - MTX-alone the poor predicted outcome in >80% cases versus
    - MTX+TNFi the for improved rate of remission up to 50+%
- and compared to routine care where patients predicted to respond based on naïve CD4+T-cells, receive MTX only (with an 80% chance to achieve remission) as standard of care.

This has been delayed by the pandemic and is still recruiting. Briefly, this is a single-centre longitudinal cohort study aiming to determine the clinical utility of naive T-cell stratification for rationalising treatment with methotrexate (MTX), for DMARD-naive early RA patients. Thus, it aims to determine whether TNFi therapy (Benepali) instituted as first-line therapy in DMARD-naive early RA patients with poor T-cell prognostication, confers better outcomes (clinical and structural). Hence, this would enable a change in practice improving outcomes for early targeted treatment for those with a poor prognosis based on their immunological status (ClinicalTrials.gov Identifier: NCT03813771) [521]. Nonetheless, a (multicentre) external validation of the value of naïve CD4+T-cell for MTX responsiveness would still be needed.

Th17 cells

No study (to my knowledge) reported on the use of flow cytometry for Th17 cells (or a surrogate phenotype of Th17 cells) to predict MTX response. However, in a study comparing early drug naïve-RA on MTX in a T2T protocol (adding sulfasalazine or hydroxychloroquine if remission was not achieved at 8-12 weeks), fewer circulating Th17 cells at baseline, quantified using a epigenetic mark on the *IL17A* gene (DNA-demethylation of a specific CpG in the *IL17A* gene promoter) rather than flow cytometry was predictive of DAS28CRP-remission after 6 months [359].

B-cells

In a study evaluating RA patients with various disease durations, treated with MXT-T2T, but adding TNFIs for those who failed to reach at least a good response based on EULAR criteria at 3 months, a higher frequency of circulating IgD+CD27- naïve B-cells at baseline was predictive of Clinical Disease Activity Index (CDAI) remission after 6-months (using OR) [522]. Patients with high circulating IgD+ CD27- naïve B cells at baseline had 4 times more chance of reaching remission than patients with lower percentages.

In new-onset RA patients, MTX + leflunomide (another scDMARD) was associated with changes in the frequency of different regulatory immune cell subsets notably decreasing Bregs (defined CD19+TIM1+IL10+ or CD19+CD5+CD1d+IL10+ B-cells) frequencies after 12 weeks, but the predictive potential of Breg was not reported [272], In steroid and scDMARD-naïve early RA (disease duration of <24 week), increased frequencies of putative Breg subset (defined as CD19+CD24^{high}CD38^{high} transitional cells) compared to health, was associated with a good EULAR response to MTX at 12 months [523].

Altogether the value of B-cell and B-subsets remains to be better investigated for predicting MTX response.

Monocytes

It was demonstrated in untreated RA patients, that responders (based on DAS28) showed steady numbers of monocytes and of their subsets ($CD14^{high}CD16-$ and $CD14^{high}CD16+$) over 6 months whereas a higher number of circulating monocytes/ subsets were potential predictors of a poor clinical response to MTX (using OR) [524]. Increase expression of CX3CR1 on monocytes was used to explain the poor predictive value of monocytes in that study, for clinical response based on a functional deficit of these cells [524]. A higher expression of this homing receptor may be a marker of refractory to treatment with MXT [524].

The percentage of $CD14^{high}$ monocytes expressing FcRIIIa/CD16 receptor at baseline in early DMARD-naïve RA patients was associated with change in DAS28 at 14-weeks post-MTX therapy and frequencies were significantly higher in EULAR non-responders compared to moderate or good responders [525].

So, in conclusion, several subsets of CD4+T-cells including naive cells, Treg and, Th17 cell, as well as Breg, and monocytes expressing CD16 were shown to have potential as biomarkers of response to methotrexate (MTX) in early RA. However, many other subsets have not yet been addressed and a comprehensive analysis allowing to address the dynamic of all lymphocyte populations together (rather than in isolation) will be needed to obtain a full picture of the value of flow cytometry here.

Table 7: Summary of candidate blood cell as biomarker for MTX treatment response in RA

cohort	study population	subsets	study design	Sample type	Drugs studied	statistical analysis	outcome measured	ref
Madrid (Spanish) ERA cohort	untreated RA patients (n=52)	1. higher monocytes and subset (CD14+highCD16- & CD14+highCD16+) 2. Increase CX3CR1 expression in Monocyte	longitudinal study: observational	Fresh PBMC	MTX	univariate analysis: modelling (t-test/ROC)	1. associated with predicting a reduced clinical response . @6months. 2. associated with non-responder	[524]
Madrid (Spanish) ERA cohort	DMARD-naïve ERA patients (n=48)	higher cTr B cells	Cross-sectional: observational	Fresh PBMC	MTX + low-dose prednisone	multiple logistic analysis : Independent association (OR)	associated with a LDA/good EULAR response @12 months	[523]
Italian RA cohort	RA (n=122)	higher naïve B cells (IgD + CD27-)	longitudinal prospective study	fresh blood	1.MTX alone 2.MTX + anti-TNF	Univariate/multivariate analysis: modelling (AUROC/ OR)	1. a significant predictor of CDAI remission @6-months	[522]
Newcastle Early Arthritis Clinic	drug-naïve early RA patients (n=68)	Higher CD39 expression on CD25 High CD4+ T-cells (putative Treg)	longitudinal prospective study	Fresh PBMC	MTX	binomial logistic regression: modelling (AUROC)	predicted DAS28CRP remission	[519]
Leeds Early Arthritis cohort	cDMARDs-naive RA patients (n=120)	higher Naive	prospective study	fresh blood	MTX	multivariate modelling (AUROC/OR)	associated with predicting MTX-induced DAS28remission	[5]
Leeds Early Arthritis cohort	drug-naïve, early RA patients (n=108)	higher naïve cell frequency	prospective study	frozen PBMC (n=38) fresh blood (n=70)	1. MTX 2.MXT+anti-TNF	Multivariate analysis (AUROC/OR)	1.associated with (DAS28<2.6) remission @6months 2. not associated with Remission	[520]

Leeds Early Arthritis cohort	Early Inflammatory Arthritis (n=28)	higher naïve T-cell frequency	pilot observational study	fresh blood	MTX alone	AUROC/OR	a predictor of remission at 6 months	[362]
Yorkshire Early Arthritis, UK cohort	DMARDs-naïve early RA (n=42)	Increased FccRIIIa/CD16 expression on CD14++ monocytes	observatory study	Fresh blood	MTX	univariate association (MWU test)	associated with EULAR non-responders	[525]
Swedish Early RA cohort	untreated early RA (eRA)(n=17).	higher PD-1+TFh higher CTLA-4+ conventional CD4+ T cells(nonTreg defined as CD25-CD127-/+)	longitudinal (randomized trial) study.	fresh PBMC	MTX+ CTLA-4Ig (abatacept)	Multivariate analysis: modelling (AUROC)	predicted CDAI remission (CDAI ≤ 2.8) @ week 24	[526]
Moscow ERA cohort	early untreated RA(n=45)	1. higher FoxP3+CD25+ 2.high level of CTLA4+ on the Treg	cross-sectional: observational study	fresh PBMC	MXT	Univariate analysis: association (MWU test)	1. associated with Good EULAR responders to MT therapy at week 24 2. associated with DAS28 remission/low RA activity @24weeks	[361]
Indian RA cohort	DMARDs naive-active RA (n=70)	1.higher CD39+ Tregs, 2. higher CD4+CD25+CD39+ cells	Cross sectional: observational	Fresh blood	MTX	univariate modelling (AUROC)	1. associated with EULAR responder group 2. associated with poor response @ 4 months.	[363]
Brazil RA cohort	RA patients (n=122)	low density of CD39 on peripheral regulatory T cells	prospective study,	fresh PBMC	MTX	univariate association (ANOVA/ t-test)	associated with non-responsiveness at least 3 months	[512]

Methotrexate (MTX), Area under (AU) receiving operating curve(ROC), Odd ration (OR), The European Alliance of Associations for Rheumatology (former European League Against Rheumatism) (EULAR), Rheumatoid arthritis (RA), Disease activity score(DAS), Clinical Disease Activity Index (CDAI),Mann-Whitney test (MWU)

1.7.3.2 Blood cell biomarkers of responses to anti-TNF treatment.

Brief summary of works focusing on subsets associated response to anti-TNF alone or with MTX treated is presented in **table 8.**

T-cells (Naive, Treg and IRC)

Naïve CD4+T-cells, Treg and IRC were investigated in TRIMID group in a small number of patients on anti-TNF over the years (unpublished abstract, presented at EWRR meeting in May 2024 [527]. This work showed that over time, naïve cell frequencies increase with anti-TNF therapy, although the full trial clinical data is currently available, and we don't know if response versus non-response will show differential patterns. In contrast, patients on MTX-alone showed continued loss of naive cells over time in non-response and a lower predictive frequency at baseline as expected, while in responders higher and steady state frequencies were observed over 48 weeks.

Similarly, Treg frequencies keep reducing with MTX-alone and remain stable with anti-TNF (with no difference at baseline). In contrast, IRC frequencies remained steady with MTX-alone but reduced with anti-TNF (again with no difference at baseline).

My group also showed in the past [491] that higher naïve cells ,lower IRC and higher Treg (with a refined phenotype adding CD45RA+and CD62L+ as newly developed Treg) provide value as a potential biomarker for the safe discontinuation of anti-TNF in early RA treated outside of NICE guidelines, in clinical trials testing the benefit of early anti-TNF while not in established RA treated according to standard of care. This was however an observational study with no predictive value assessed. It would need repeating with similar patients

from biologics trials in early RA, which was an objective of my studies, that will not be achieved due to the pandemic.

Th17 cells

A lower Th17 cell frequency (based on CD45RA−CD161+CCR6+ markers) was shown to be associated with a good EULAR response at 6 months in active RA patients treated with anti-TNF as well as anti IL6 (tocilizumab) [528]. But with small number of participants (n=31), the statistical power to suggest Th17 robust association with responses may be lacking. Additionally, the gating strategy used to identify Th17 cells was unusual. T helper type 17 (Th17) cells are a subset of $CD4^+$ effector T-cells [529]. The author defined Th17 as part of the CD45RA- subpopulations (the memory compartment) of the total CD4+T.

The study based this gating strategy on an animal study (using a C57/BL6 mice model) [530] which showed that upregulation of IL-17 exclusively occurs within the memory (defined by $CD44^{high}$ $CD62L^{low}$) compartment of T-cells after activation to induce Th17 differentiation, hence not in naïve ($CD44^{low}$ or $CD62L^{high}$) cells compartment.

On the other hand, Th17 differentiation from naïve CD4+ T lymphocytes has also been reported in another animal studies [531] and in human [532, 533]. Applying evidence from animal studies to define Th17 cell may also be misleading as CD4+T-cell are likely to behave differently in humans. Therefore, quantifying Th17 only in the memory compartment (CD62L+ CD45RA- cells) based on the co-expression of CD161+CCR6+ (a 2-maker identification alone, not the production of IL-17) [528] probably may have biased the data. Using a more definitive surrogate phenotype notably, 3-makers (CD161+ CCR6+ CXCR3- CD4+ T-cells) for th17 [359] may have been more suitable while not yet optimal.

The percentage of CD8+T-cells at BL, specific for antigens derived from apoptotic cells in chronic inflammation was significantly higher in ant-TNF responders compared to those who failed to respond following treatment [534], this was decreased, suggesting a possible involvement of apoptosis epitope-specific CD8+ T cells in pathways relating to the mechanism of action of anti-TNF. While high CD8+T presented a potential predictive biomarker of EULAR response after 6 months of anti-TNF (AUC 0.82, n=16) in this single study, there would be a need to validate the finding using a larger cohort independently to assess it potential biomarker value.

B-cells

B cells have been identified as a potential predictive biomarker of response to anti-TNF [535] with data suggesting that a higher frequency CD27+ memory B-cells, notably IgD+CD27+ pre-switch memory B-cells, have value as a biomarker for predicting EULAR responders at 3 months based on EULAR criteria (using ROC, relative risk).

A report from UK cross-sectional cohorts plus European prospective cohorts (part of the ABIRISK consortium) had shown that a reduced frequency of memory B-cells (defined by the limited phenotype $CD19^{+}CD24^{hi}CD38^{lo}$) expressing a signal regulatory protein (SIRP a/b) predicted patients that developed anti-drug antibodies (ADA) which was associated (using *t*-test analysis) with non-response or partial response to anti-TNF (adalimumab) treatment according to the EULAR classification at 12 months with AUROC score of 0.92 [536].

About 33% of adalimumab-treated RA patients developed immunogenicity. The development of ADA in RA patients treated with biologics has been shown to vary depending on the drug in use [537]. This is because b-DMARDs are proteins

which are intrinsically immunogenic as foreign (such as neutralizing anti-drug antibodies, contributing to loss of efficacy) [538-545] .Therefore, the development of ADA may not only impact drug efficacy but also elicit adverse drug reactions. As such, it was argued that for patients who fail an initial biologic agent or for those who initially respond but lose efficacy, the evaluation of whether ADA is present will allow clinicians to more effectively choose the best agent to which the patient can be switched [537] as such, reduced frequency of memory B cells ($CD19^{+}CD24^{hi}CD38^{lo}$) particularly those expressing this signal regulatory protein (SIRP a/b) was proposed to be a good predictor (using AUROC) for the development of ADA and by extension a potential predictive biomarker for poor clinical response to adalimumab.

In a more recent study Rodríguez-Martín et al., [546] phenotyped multiple blood lymphocyte subsets in bDMARDs naïve-RA patients treated with anti-TNF (after MTX failure to achieve remission). The data seems to validate the work done by Fedele et al., [522] as lower % of total and naïve B cells at baseline were found to be associated with DAS28 and SDAI non-remission at 6 months. The study further demonstrated using a multivariate logistics regression, that a B-cells/CD4 ratio <0.2 was highly predictive of DAS28 non-remission after 6 months. However, this data was not replicated in a more recent study [371] which demonstrated that lower percentage of naive B-cells was independently associated with remission based on DAS28 in anti-TNF +MTX treated patients at 6 months. No association, however, was observed for patients who did not receive concomitant MTX in the study suggesting that the variance between findings could be a consequence of a combination therapy in some studies.

Monocytes

A study of MTX non-responders RA patients treated with anti-TNF (adalimumab) showed that good clinical response was found in patients with a significant normalization of the numbers of circulating monocytes ($CD14^{+high}CD16^{-}$ & $CD14^{+high}CD16^{+}$) after three months of treatment which then lasts up to six months. Although at baseline, the cellular markers were not showing a predictive value for discriminating between responders and non-responders, the study demonstrated a high positive predictive value (86%) using numbers of monocytes at 3 months as an early biomarker for anti-TNF response which may be of help to manage people with no response by switching to other drugs early [547]. This is univariate modelling using only monocyte without accounting for the effect of clinical variables in the model (using AUROC). Therefore there is a need to validate this finding in an independent cohort while accounting for other possible confounders.

In conclusion, multi-centre replication studies would be needed to validate the clinical value of these potential cellular biomarkers to better inform the choice of using anti-TNF or otherwise in patients.

Table 8 Summary of candidate blood cell as biomarker for anti-TNF treatment response in RA

cohort	study population	subsets	study design	Sample type	Drugs studied	statistical analysis	outcome measured	ref
Madrid RA cohort	anti-TNF naïve-RA (n=78)	lower naïve B-cell	a prospective, observational, longitudinal bi-center pilot study	frozen PBMCs	1.anti-TNF +MTX 2.anti-TNF alone	Multivariate logistic analysis: modelling (OR/ AUROC)	1. independently associated @ 6 months. 2. no association	[371]
Madrid (Spanish) ERA cohort	RA patients (n=98): Biological naive(n =86) + TNFi-treated(n=12)	1. lower total and naive-B 2. B-cell/CD4 ratio<0.2	prospective observational, longitudinal, bi-center pilot study	frozen PBMCs	anti-TNFs	Univariate (MWU test) and multivariate analysis: modeling (OR/AUROC)	1. associated with no REM @ 6months. 2. associated with a higher probability of non-REM status	[546]
Madrid (Spanish) ERA cohort	MTX non-responder RA patients(n=35), active drug naive-RA(n=13)	1. higher monocytes (CD14+highCD16- ,CD14+highCD16+ and CD14+lowCD16+ subsets) 2. Increase CX3CR1 expression in monocyte	longitudinal study: observational	Fresh PBMC	anti-TNFa (adalimumab) + MTX	univariate analysis modelling (t-test/ROC)	1. @ 3months of treatment associated with Non-responder @6months. 2. associated with non-responders	[547]
French RA cohort	RA(n=20)	higher CD27+ memory B cells (IgD+CD27+ pre-switch memory B cells)	pilot observational study		ant-TNFs	Univariate analysis: association (MWU test)	associated with EULAR responders @3 months.	[548]
Roma, Italian cohort	HLA-A2+ biologic-naïve RA(n=16)	High apoptotic epitope-specific CD8+ T cells		PBMC	anti-TNF (Etanercept)	univariate linear regression: prediction (AUC)	associated with EULAR response@6onths	[534]

Leeds, UK cohort	RA remission (n=47)	lower IRC higher naïve higher CD62L+Tregs	prospective controlled study	whole blood	anti-TNF+MTX (accesation)	logistical regression: univariate associations (AUROC/OR)	associated with Sustained remission was associated	[491]
London, UK cohort + European (France, Netherlands and Italy) cohorts	RA (n=57)	reduced frequency of signal regulatory protein (SIRP)a/b-expressing memory B cells	UK cross-sectional cohort(n=20) + European prospective cohort(n=37)	frozen PBMCs	anti-TNF (adalimumab)	univariate analysis: modelling (AUROC)	predicts development of ADA, and consequentially non reponders@12months	[536]
France RA cohort	RA(n=96)	higher proportions of CD27+ memory B cells	longitudinal prospective study	PBMC from EDTA blood	anti-TNF	Multivariate linear regression analysis: model (AUROC)	associated with responders (according to EULAR) at 3 months	[535]

Anti-Tissue necrotic factor(TNF), Methotrexate (MTX), Area under (AU) receiving operating curve(ROC), Odd ration (OR), The European Alliance of Associations for Rheumatology (former European League Against Rheumatism) (EULAR), Rheumatoid arthritis (RA), Mann-Whitney test (MWU), remission (REM), peripheral blood mononuclear cell (PBMC)

1.7.3.3 Blood cell biomarkers of responses to CTLA4-Ig (abatacept) treatment

CTLA4-Ig, commonly known as abatacept, targets the immune system by inhibiting the co-stimulatory signal required for full T-cell activation. Specifically, it binds to CD80 and CD86 on antigen-presenting cells (APCs), thereby preventing their interaction with CD28 on T cells. This interaction blockade inhibits the co-stimulatory signal necessary for T-cell activation, ultimately leading to suppression of the inflammatory response involved in autoimmune diseases like rheumatoid arthritis (RA) [549] . While abatacept has demonstrated efficacy in many patients, response rates vary, prompting the need for predictive biomarkers to identify individuals most likely to benefit from treatment [550, 551]. Blood cell biomarkers represent a promising avenue for predicting responses to abatacept therapy, as they reflect the dynamic interplay between the immune system and disease activity [551]. This section explored the current understanding of blood cell biomarkers associated with responses to abatacept treatment in RA, highlighting their potential utility in personalized medicine approaches and guiding treatment decisions for improved patient outcomes.

T-cells subsets

In early RA patients, higher frequencies of activated Treg ($CD3^+$ $CD4^+$ $CD25^+$ $CD127^{low}$ HLA-DR+) were selected at baseline in a model of remission-related biomarkers based on clinical disease activity index (CDAI) after 6month abatacept, adjusting for other parameters [552]. In established RA, activated Treg were not found to be associated with remission [552] suggesting that Treg dysregulation may play a critical role in the onset of RA.

In a clearly defined cohort of drug-naïve early RA patients, Aldridge et al., [526] investigated $CD4^+$ T-cell subsets and found that higher proportions of PD-1+TFh and CTLA-4+ conventional CD4+ T cells at baseline (using AUROC) predicted remission (CDAI ≤ 2.8) at week 24 in patients treated with abatacept (CTLA-4Ig).

Patients with low baseline numbers of $CD8^+CD28^-$ T cells showed a greater than 4-fold higher probability (odds ratio) of achieving remission within 6 months of abatacept therapy than patients with higher levels of these cells [553]. This suggested that baseline frequency of circulating CD28-negative CD8+T cells may discriminate between potential responders and non-responders to abatacept in patients with RA although, it was not demonstrated using appropriate statistical methods by the authors.

Th17 cells

Additionally, activated Th17 ($CD3^+$ $CD4^+$ $CXCR3^-$ $CCR6^+$ $CD161^+$ HLA-DR^+) cells at baseline were shown to be significantly higher in remission versus non-remission and were selected using a Cox regression model-related method as a remission-related change in T-cells features in the early RA treated with a biologic (abatacept). The remission rate however varied between RA subtypes, the proportion of patients achieving remission being lower in the seronegative compared to seropositive patients. Th17 cell subset was therefore presented as a predictive biomarker for good clinical response to abatacept in patients with early, seropositive rheumatoid arthritis [552].

B-cells

Lower baseline levels of activated B cells (IgD+ CD38+) and/or memory B-cells (CD27+) (measured as absolute count) were shown to be important for an abatacept poor clinical response in RA resistant to other biologics (anti-TNFα, anti-CD20, RTX and anti-IL6R, Tocilizumab) based on DAS28 after 6 months [554]. Salomon et al, [528] found in active RA patients who required the initiation (or switch) to a biologic drug, that baseline proportion of $CD24^{hi}CD27^{+}$ Breg was associated (MWU test) with DAS28 remission at 6-months and that abatacept-treated patients with good EULAR response at 6-months had significantly higher proportion of Breg cells than those without good EULAR response, however, only using univariate analysis with no further evaluation of predictive value of the potential biomarker. **Table 9** presented a brief summary of the candidate subsets report to be associated with response to CTALA4-Ig therapy.

Taken together, investigating the value of other immune cells as biomarkers of clinical response to co-stimulation blockade (CTLA4-Ig) treatment is needed in the future.

Table 9 Summary of candidate blood cell as biomarker for CTALA4-Ig (abatacept) treatment response in RA

cohort	study population	subset	study design	sample type	statistical analysis	outcome measured	ref
Brest, France RA cohort	RA resistant to other biologics (n=43)	reduced CD38+ and/or CD27+ memory B cell	a retrospective monocentric study	fresh blood	univariate association (MWU test/ Wilcoxon test)	poor clinical response (NR/MR) based on DAS28 @ 6months	[554]
Brescia, Italy RA cohort	RA anti-TNF-resistant RA(n=32)	Low CD28− CD8+ T cells		whole blood	univariate prediction (AUROC)	associated with predicting remission based on EULAR response criteria after 6 months	[553]
Tokyo RA cohort	RA(n=103): 1. seropositive early (n=24) *2.* seropositive established (n=79*)*	1. higher activated Th17 (aTh17) 2. higher activated Treg (aTreg)	longitudinal study: observational	PBMC	1. Cox regression model–related method (AUROC) 2.Univariate logistic analysis: association	1. predictors of remission based on CDAI @ 6 months 2. associated with a good response	[552]

Area under (AU) receiving operating curve(ROC), The European Alliance of Associations for Rheumatology (former European League Against Rheumatism) (EULAR), Rheumatoid arthritis (RA), Mann-Whitney test (MWU), no response (NM), moderate response(MR), Disease Activity Index (CDAI), peripheral blood mononuclear cell (PBMC)

1.7.3.4 Blood cell biomarkers of responses to B cell depletion (Rituximab) treatment

Rituximab (RTX) is a B-cell depleting antibody therapy that only affect CD20+ B-cells leading to their depletion from the peripheral circulation and lymphoid tissues. Notably, rituximab does not affect plasma cells, which lack CD20 expression, thus sparing the humoral immune response mediated by long-lived plasma cells. It showed good clinical response in RA[555, 556] but is most effective in systemic lupus erythematosus (SLE)[557, 558] . The rational for investigating blood cell biomarkers associated with RTX response in RA lies in the complex interplay between B-cells, T cells, and cytokines in the pathogenesis of the disease[559] . B-cells play a central role in RA pathophysiology by producing autoantibodies, presenting antigens to T-cells, and secreting pro-inflammatory cytokines. Targeting B-cells with RTX can modulate these immune responses, leading to clinical improvement in some patients. Several studies[555, 556] have suggested that baseline levels of specific blood cell subsets, such as CD20+ B-cells, and CD19+ B-cells, may predict response to RTX therapy in RA.

<u>B-cell subsets</u>

Higher total lymphocyte counts (LC), and higher plasmablasts frequency using ROC analysis were independently shown to be predictors of poor response to RTX treatment as well as failure to achieve remission at 6 months (sensitivity [93.3%]/specificity [44.8]) [560]. The success of RTX monotherapy after 6 months was reported by our team to be associated with complete B-cell depletion after 1st infusion in both biologic naive RA and patients who were resistant to anti-TNF treatment [561]. Higher plasmablasts following incomplete B-cell depletion, have been observed in poor clinical response [561]. The 2nd cycle of rituximab (administered prior to total B cell repopulation) enhanced B cell depletion and

clinical responses, and further strengthened that clinical response was mediated by the extent of B cell depletion rather than the RTX dose.

A moderate-to-good response was reported after 6 months in a small number of RA patients (resistant to DMARDs including at least one anti-TNF agent) who switched to rituximab treatment [562]. This clinical response was preceded by a more significant decrease in peripheral blood $CD19^{+}CD27^{+}$memory B-cells, suggesting that better depletion of memory B-cells was associated with treatment response using t-test analysis [562].

Baseline levels of activated switched memory (SM) CD95+ and CD21- (IgD-CD27+) and double negative (DN) memory (IgD-CD27-) B cells were found not to be predictive of clinical response [563]. However, the study further suggested that only lower CD95+ activated memory B cells (CD95+IgD-CD27+ and CD95+IgD-CD27-) at depletion time points (1-month post BCDT) have the potential for biomarkers of clinical response after 4 months of follow-up using MWU test/Kruskal-Wallis multiple comparison tests.

A 2-year national, multicenter, randomized, open-label RTX retreatment study, demonstrated in BCDT–naive RA patients, that a low baseline CD27+ memory B cell frequency was predictive of a greater clinical response to RTX (odds ratio 0.97) 6 months after the first cycle based on DAS28-CRP [367].

A combination of B-cell numbers (CD19+ B cells at a level greater than the lower limit) with increased memory B-cells (CD19+CD27–IgD–) was demonstrated as independent predictors of response (OR=2.2 [1.4, 3.5]) in patients failing one anti-TNF [564]. Looking at a combination of both B-cell biomarkers at baseline, better treatment effects were noted particularly in RF+ RA based on DAS28, whereas

only a higher proportion of memory B cells identified responder within RF-patients [564].

In RA patients resistant to TNF-blocker, a significantly higher frequency of naïve B-cells (CD27-IgD+) at baseline was reported in non-responders/moderate responders to RTX compared to good responders [365]. A low frequency of plasmablasts pre-treatment was also proposed as a valid predictor for EULAR responsiveness in TNF-resistant active RA patients [366] as naïve/memory B-cell frequencies are in apposition; this suggests that the markers/phenotype used may have an influence on the results by selecting different subsets with better statistical value than others and shows that a careful selection of surface marker may have a major effect on findings.

NKT/NK cells

In a very small study (n=7), lower numbers of circulating iNKT cells (defined by CD3, CD4, and CD161) in RA patients at baseline [565] were shown to be significantly increased in responders (n=5) after 120 days of rituximab treatment, and inversely correlated with DAS28, which suggested a potential predictive role.

Data have shown that NK cells(CD3−CD56+) could be used as an early predictor of clinical response, the study demonstrating increased activation of NK-cells defined as an upregulation of CD54 (CD56+/CD16+/CD54bright) 3 months after the first rituximab course, which was significantly associated with clinical response (using linear regression model analysis) at 6 months and 1 year, independent from other clinical variables [566].

T-cells

On the other hand, lower total LC and CD4+T-cells at baseline were independently associated with a good EULAR response to RTX in RA despite RTX targeting B-cells. However, these parameters failed to predict patients who would achieve remission. Interestingly, unpublished data from TRIMID group (presented at EWRR conference, 2019) showed that the return of T-cells to the synovial biopsies was associated with relapse after RTX therapy suggesting that there may also be differences between the drug target (B-cells) and the cellular biomarker (T-cells) that predict the response.

In summary, B-cells seemed mostly investigated **(Table 10)** as a biomarker of RTX response, but it may not be the only lymphocyte subset (LS) with value. Therefore, highlighting the need to investigate the value of other LS as a biomarker of RTX response

Table 10 Summary of candidate blood cell as biomarker for RTX treatment response in RA

cohort	study population	subset	study design	sample type	Drugs studied	statistical analysis	outcome measured	ref
Strasbourg, France cohort	active RA(n=7)	lower iNKT	longitudinal analysis	fresh blood	RTX	univariate associations (MWU test)	potentially associated with good clinical responses @120days	[565]
Italian RA cohort	long standing RA (LSRA) (n=138)	lower lymphocyte count	retrospective observational (Muti-centre) study		RTX	univariate logistic regression analysis/multivariate: Modelling (OR)	predictors for good-EULAR response	[369]
Magenta, Italy cohort	RA refractory to conventional and anti-TNF (n=34)	increased activated NK count (defined as an upregulation of CD54)(CD56+/CD16+/CD54bright)	longitudinal study: observational	fresh blood	RTX	Multivariate linear regression "least squares" models	significant inverted relationship between SDAI/DAS28 @6months	[566]
Leeds Early Arthritis Cohort	RA(n=19)	complete B-cell depletion after 1st infusion	pilot study	Fresh blood	RTX	univariate association (MWU test)	associated with good clinical response @ 6 months,	[561]
Leeds Early Arthritis Cohort	RA(n=103)	higher pre plasma + incomplete depletion	prospective study	Fresh blood	RTX	univariate logistic regression: association	associated with non-responders	[567]
Leeds Early Arthritis Cohort	anti-TNF unresponsive active RA(n=60)	complete B-cell depletion after 1st infusion	prospective study	Fresh blood	RTX	univariate association (MWU test)	associated with moderate-to-good clinical response @ 6 months,	[568]
London RA Cohort	active refractory RA(n=24)	higher numbers of memory B cells @ rtp	prospective observational study.	fresh blood	RTX	Wilcoxon test	associated with RA relapsed	[569]
New York RA cohort	RA active RA(n=20)	lower CD95+ activated memory B cells (SM and DN) at dtp higher ratio of transitional B cells to memory at rtp		fresh PBMC	RTX	Univariate analysis: association (MWU test)	Biomarkers of clinical response @ 4 months	[563]

Graz, Austria RA cohort	anti-TNF treated active RA(n=52)	1 elevated CD95+ pre-switch B cells 2.lower Plasmablasts	Cross-sectional: observational	PBMC	RTX	univariate logistic regression analysis (MWU test/ OR)	1.associated with non-responders 2. predictor for EULAR responsiveness	[366]
Graz, Austrian RA cohort	RA(n=44)	1.lower total lymphocyte counts (LC), T cells and CD4 + T cells 2.(high LC (absolute value), high plasmablast)	observatory study	Fresh blood	RTX	Multivariate logistic regression: modelling (AUROC)	1. independent predictors of EULAR response 2. predicted not achieving LDA at week 24	[560]
Heraklion, Greece cohort	Active RA resistant to DMARDs including at one anti-TNF agent (n=31)	lower CD19+CD27+ memory B cells	Cross-sectional: observational	fresh PBMC and BM	RTX	univariate association (Wilcoxon test)	associated with moderate-to-good response (n=6) @6month	[562]
Bern, Switzerland RA cohort	RTX-naïve' RA resistant to TNF-blocker(n=35)	lower naïve (CD27-IgD+) B-cells (absolute number)	prospective observational study.	fresh PBMC	RTX	univariate association (Wilcoxon test/ MWU test)	associated with non-responders or moderate responders	[365]
France RA cohort	BCDT–naive RA patients (n=208)	low CD27+ memory B cell frequency	RTX retreatment study, A 2yrs national, multicenter, randomized, open-label study	fresh PBMC	RTX	Univariate analysis: modeling (OR)	associated with a greater clinical response 6MONTHS	[367]
Germany RA cohort	RA patients failing one TNFi(n=154)	increased CD19+CD27–IgD– B cells (absolute value)	exploratory, multicentre, open label, uncontrolled phase IIIb studies	Fresh blood	RTX	multivariate regression: modelling (AUORC/OR)	associated with responder group after 16 weeks	[564]

Rituximab (RTX), Area under (AU) receiving operating curve(ROC), Odd ration (OR), The European Alliance of Associations for Rheumatology (former European League Against Rheumatism) (EULAR), Rheumatoid arthritis (RA), Mann-Whitney test (MWU), remission (REM), Depletion time point(dpt), reconsitution time point(rtp), rituximab (RTX), low disease activity(LDA), Simplified Disease Activity Index (SDAI), Disease Activity Index (CDAI), peripheral blood mononuclear cell (PBMC): Blood cell biomarkers of responses to Tocilizumab (IL-6 inhibition) treatment

Despite advancements in treatment modalities, achieving optimal outcomes for all patients remains a challenge, necessitating a personalized approach to therapy. Among the arsenal of biologic agents, tocilizumab, a monoclonal antibody targeting the interleukin-6 (IL-6) receptor, has emerged as a pivotal player, offering promising results in modulating the immune dysregulation inherent in RA. However, the heterogeneous nature of RA poses a conundrum for clinicians, as responses to tocilizumab treatment vary widely among individuals[509, 570] . Hence, the quest for reliable biomarkers capable of predicting and monitoring response to therapy has intensified. In this pursuit, blood cell biomarkers have garnered significant attention due to their accessibility, dynamic nature, and potential to reflect the underlying immunological milieu[570-573].

There are fewer studies here as this drug was more recently developed. Investigation of lymphocyte lineage subsets comprising T, B, NK and NKT cell for response to tocilizumab in patients with severe rheumatoid arthritis (n=20) observed that only NK-cell (CD3−CD56+ cells) in higher proportion at baseline was associated with disease remission at 3 months (using MWU test) [484]. The failure of the baseline proportion of NK cells to show association with change in disease activity at 3 months in a second cohort (n=15) treated with anti-TNF therapy was suggesting that NK cells may be a biomarker of response specific to an IL-6 inhibitor agent with a pathological link between anti-IL-6 induced remission and NK cells.

Conclusion to this section on response to drug

Despite the substantial increasing evidence of various cellular signatures underlying RA pathogenesis, a comprehensive lymphocyte analysis of the dynamics of all subsets together for predicting clinical response in RA has not been done. It remains important to predict more accurately which patients are more likely to respond to a targeted treatment to improve the risk-benefit ratio and cost-effectiveness in individual patients as well as the overall treatment success on the population level [574, 575].

1.7.4 : In summary

From the current review, the lack of results validation in independent cohorts has been found to be the most important limitation of many studies across each stage of the IAC. While some results have been replicated in similar cohorts by the same group as ours, others have not. This highlights the gap in translating promising predictive biomarkers reported by most studies into routine clinical tests. A need that seeks urgent attention in this field of study. Although many studies reviewed in this section had proposed promising predictive biomarkers across various stages of the IAC, most were significantly limited by a few factors, one such factor was;

The size of the study population.

Only 14 out of 55 studies included >100 participants, 14 more enrolled < 100 subjects, while 27 studies had less than 50 subjects **(Table 5 to Table 10)**. With a smaller sample size, it is challenging to draw a definitive conclusion on statistically significant associations because of the possibility of higher standard error arising from a lack of statistical power.

The validity of the study

Another concern lies in the validity of the statistical model reported by some studies. For clinical decision-making about patients, diagnostic and prognostic inferences from statistical models are crucial. How reliable these inferences are is dependent on the statistical model's validity as well as the accuracy and completeness of the data gathered [576]. Prediction models reported by some studies lack the appropriate statistical rigor underpinning the value of the model as claimed.

While some studies used multivariate analysis to reach model construction, others demonstrated only univariate association which failed to adjust for other possible confounders and used inappropriate wording to describe findings (i.e., prediction of which using association statistics).

The end point for outcome assessment was another factor that made comparison of findings from different studies difficult. It is not always clear if biomarkers are associated or more importantly predictive of those achieving the outcome of interest while the time at which the outcome is assessed is also quite valuable. Potentially, stable response for an impact on a clinical decision may not be accounted for in studies accessing biomarkers at a shorter treatment duration (like at 3 or 4 months) [363, 484, 535, 548, 562, 563, 565] or in patients, 3 months after treatment and not at BL [512, 547].This may be responsible for inconsistency in data reported in literatures evaluating baseline cell subsets that might influence clinical response to therapy used may also bring bias in the data.

Non-uniformity in classifying response types

Additionally, non-uniformity in classifying response types may partly explain data inconsistency. While some studies [525, 554] used 3 categories (good, moderate or non-response) according to EULAR criteria [577, 578], other have used 2 types- either as moderate and good responders [369] or responder (combining good + moderate) and non-responders [525, 534, 547] or good responders and poor responders (combining moderate + non-responders) [523, 528] or remission and non-remission [5, 371, 522, 528, 546]. Does (good)responders mean those in remission? or does (moderate)responders mean non-remission? These are conceptual questions that needs to be clearly defined to enable better data comparison across studies. Agreement between multiple response criteria may be poor and will likely cause data variability between studies on similar biomarkers, therefore, this needs to be taken into consideration drawing conclusions when on reproducibility and validity of biomarkers.

Different techniques of analysis/variability in defining cell phenotype

There are also several pieces of literature quantifying lymphocyte subsets with major inconsistencies attributable to the use of different techniques, as well as the use of different markers to describe the same cell “phenotype”. Th17 [579, 580], Treg [5, 361, 363, 581] and Breg [270, 582, 583] have widely been reported with varying phenotypes.

While some studies only evaluated the efficacy of a single treatment, they often included patients who had received a range of treatments, sometimes targeting various molecular pathways such as anti-TNF and IL-6 inhibition prior to the study BL. It is highly probable that a biomarker is predictive of response to one specific

class of therapy but not another because of the immune mechanism it is targeting. However, most of these targets are part of pathways that are not working in isolation but in a network/relationship with each other. Hence, interpreting data where patients are treated with various classes of agents is difficult and it would require clear stratification of where patients are in their disease/drug history to ensure a biomarker is valid as well as not discarded because not for the right group. Therefore, there is a need for better characterisation of the patient's history (age, smoking status, gender, first degree relative, duration of disease) for improved personalised medicine to be achieved in the management of inflammatory arthritis such as RA.

1.8 Predicting treatment outcome in Psoriatic arthritis (PsA)

PsA is a chronic, systemic inflammatory arthritis often linked to the presence of skin psoriasis (Pso). It ranks among the most prevalent inflammatory arthritides, affecting up to 0.1% of the general population and manifesting in 6% to 42% of individuals with psoriasis [584, 585]. It manifests as a heterogeneous spectrum of musculoskeletal and dermatological manifestations, often accompanied by enthesitis, dactylitis, and nail dystrophy. PsA significantly impacts patients' quality of life due to its chronic nature, potential for joint damage, and associated comorbidities [586, 587].

While the exact aetiology of PsA remains elusive, it is believed to arise from a complex interplay of genetic predisposition, environmental triggers, and dysregulated immune responses. Notably, the pathophysiology of PsA involves aberrant cytokine signalling, particularly TNF-α, IL-17, and IL-23, which drive inflammation, synovial hyperplasia, and tissue damage [586-588] .

Over the past two decades, accumulating evidence in RA has underscored the significance of early and aggressive treatment with DMARDs and biologics in mitigating the adverse effects of the disease, particularly in terms of joint damage and disability [589] .

This paradigm shift has transformed clinical practice, with clinical remission becoming an attainable goal for many patients, especially those initiating biologics, predominantly TNF blockers, early in the disease course [514, 590]. However, the landscape of evidence in PsA remains relatively sparse, with limited data on the efficacy of conventional DMARDs such as methotrexate or leflunomide [591] . In recent years, advancements in understanding PsA pathogenesis and therapeutic strategies have revolutionized its management. Biologic agents targeting TNF-α, IL-17, and IL-23 pathways have emerged as cornerstone therapies, offering substantial efficacy in mitigating disease activity, halting structural damage, and improving patient outcomes [588, 591, 592]. Thus far, TNF inhibitors exhibit efficacy across the heterogeneous clinical manifestations of the disease, encompassing joints, spine, skin, and nails [593-595] . Mirroring their impact in RA, TNF inhibitors have also demonstrated the ability to arrest structural damage progression in peripheral joints in PsA[596] . Moreover, recent advancements have introduced novel treatment modalities targeting TNF inhibition, such as golimumab, which have exhibited favourable efficacy in managing joint and skin manifestations of PsA, particularly in cases resistant to conventional DMARDs such as MXT [597] .

PsA poses a significant therapeutic challenge due to its heterogeneity in clinical presentation and variable treatment responses among individuals [588, 598-601]. While considerable progress has been made in the development of targeted therapies, identifying optimal treatment strategies for individual patients remains

elusive. Therefore, there is a pressing need to advance precision medicine approaches by harnessing the predictive power of cellular biomarkers. Cell biomarkers offer a promising avenue for predicting treatment outcomes in PsA by providing insights into the underlying pathophysiological mechanisms and patient-specific disease characteristics. This holds the potential to serve as valuable prognostic biomarkers and guide individualised therapeutic decision-making.

1.9 Flow cytometry as a biomarker technology

Flow cytometry (FC) is a technique used for the rapid extraction of multiparametric qualitative and quantitative data from individual cells within complex populations of cells suspended in a fluid [602]. It has increasingly been recognised as a highly powerful and versatile technology both in research and hospital settings not only for immunophenotyping of a variety of cells (and their subsets) but also for assessing membrane-bound and intracellular proteins, cytokine expression, or DNA contents of cells in (patho)physiological states [602-604]. As such, flow cytometry can simultaneously provide information about the phenotypic and functional characteristics of cells in heterogeneous body fluids such as blood, urine, and bone marrow, as well as in enzyme-digested solid tissues, thus enabling quantification of large numbers of cells and evaluation of the distribution of various characteristics at population levels [605]. The greatest benefit of multi-colour flow cytometry is its high specificity for discrete cell subsets and rare populations. It is indeed possible to detect rare cell subsets population with frequencies as low as 0.01% [605, 606].

Overall, FC is

- Highly effective, high throughput system for the characterization of diverse cell populations simultaneously. Hence representative even for small population of cells.
- Sensitive technique for gathering data on multiple parameters such as the expression of surface markers or the presence of intracellular cytokines and proteins with limited technical manipulation during sample acquisition not too time-consuming [607].

1.9.1 FC biomarker discovery and clinical application

FC is one of the most successful single-cell analytical tools notably utilised to characterise immune cells [608-611]. FC has an extremely wide field of application (Table 1 and 2) [612-615]. It can perform different types of analytics including immunophenotyping (extra- and intra-cellular staining) cell proliferation versus cell death, and specific protein evaluation (such as for phosphorylation) and more recently hybridisation to nucleic acids [610, 611].

Accumulated reports have demonstrated the value of FC as a technology to quantify biomarkers in diagnosis or for predicting and monitoring several disease outcomes. For instance, enumeration of CD4+T cells remains the best method of choice for monitoring the immune competence in HIV patients, which allows for better treatment management [616]. The significance of monitoring CD4+Treg cells as a prognostic biomarker in several cancers including breast, gastrointestinal, lung, and ovarian carcinoma has been documented [617]. Enumerating CD16+ NKT-like cells in colorectal cancer patients was independently associated with shorter disease-free survival [618].

In the field of autoimmune inflammatory arthritis, growing evidence in RA, systemic lupus erythematosus (SLE), Osteoarthritis (OA), Psoriatic arthritis (PsA) have demonstrated that monitoring of certain immune cells.

For instance;

- CD4+T-cell and subsets (naïve, memory, regulatory T-cells)
- B-cells and subsets (naïve, memory and regulatory B-cell)
- Natural Killer (NK CD56) cells,
- NKT-cells and
- Monocytes,

showed potential biomarker value and were considered promising with respect to diagnosis or predicting disease progression and/or treatment responses in patients [1, 3, 5, 7, 9].

The role and significance of CD4+T subsets quantification in RA patients (notably loss of naïve- and regulatory-CD4+T cells and the appearance of inflammatory related-naïve and memory cells (IRC) has been reported and furthermore replicated/validated [1, 619, 620]. The specific value of these 3 CD4+T-cell subsets biomarkers has been demonstrated by independent studies for the prediction of clinical outcomes such as flare following c-DMARDs-induced remission [1], sustained remission following cessation of anti-TNF therapy [621] and in early RA for predicting patients with a better chance for achieving remission with MTX [5, 616] as well as for progression to active disease [3].

Altogether, FC offers the best approach to safely achieve rapid quantification of peripheral blood cells for the management of diseases and has the potential to become a useful tool in most clinical and research settings to predict/monitor responses to therapies [616] as highlighted by the many applications used in clinical practice in **Table 11.**

Table 11: Flow cytometry-based assay currently in use for clinical application

Clinical application	Ref
Hematologic disease and/or Malignancies: • Minimal Residual Disease (MRD) • Lymphocyte Subset Enumeration for Immunodeficiency Disease • Reticulocyte Enumeration for diagnosis and monitoring of anaemia and bone marrow regenerative activity after chemotherapy or bone marrow transplantation • Platelet Function Analysis	[622-632]
Analysis of DNA Ploidy or proliferation and Cell Cycle	[633-638]
Measurement of the Efficacy of Cancer Chemotherapy	[639-642]
Cell function Analysis: • Protein phosphorylation • Intracellular calcium flux • lymphocyte activation (in viral infection and other diseases for example) • Oxidative burst in neutrophil as a screening test for chronic granulomatous disease • Lymphocytes neoantigen expression	[643-646]
Transfusion Medicine: • Detection of feto-maternal haemorrhage • study of HbF levels • frequency of adult red cells with low levels of HbF in individuals with hemoglobinopathies, • the medical evaluation of anaemic patients, including sickle cell and thalassemic patients	[647-655]
Organ Transplantation and Hematopoietic Cell Therapy: • Pre-transplant cross- matching (in bone marrow transplantation) • HLA antibody screening, • post-transplantation antibody monitoring. • pre-transplantation determinations of the efficacy of ex vivo T-cell graft depletion, • post-transplantation evaluation immune recovery ○ graft rejection, graft-versus host disease, graft-versus-leukaemia effect	[656-662]
Microbiology: • Microbe detection (bacteria, fungi, parasites and viruses) • quantitative procedures to assess antimicrobial susceptibility and drug cytotoxicity	[612, 663-667]

In addition to its potential for clinical management, FC has shown wide application in research. For instance, in studies involving biology of diseases such as cancer to the therapeutic and prognostic relevance of these markers (with a few studies summarised in **Table 12.**

Table 12 Flow cytometry-based application in research

Disease type	Biomarkers	ref
Ovarian carcinoma	Adhesion molecules: Integrins and the immunoglobulin superfamily member CD44 chemokine receptor: CXCR4 / CXCL12 leukocyte classes and chemokine receptors: NK cells, CD19-positive cells, CXCR4, CCR5, and CCR7	[668-670] [671] [672]
gastric carcinoma	CD44 expression and epidermal growth fac-tor receptor (EGFR) Protease: matrix metalloproteinase (MMP)	[673] [674, 675]
breast carcinoma	c-ErbB-2, member of the EGFR family	[676]
Autoimmune inflammatory arthritis	Naïve CD4+T and Treg And other subsets	[4, 5, 7, 9, 312, 616]

1.9.2 Challenges in the use of flow cytometry technology

Technically, FC workflow is not particularly challenging. Most protocols use standard procedures, optimised through years of experience both in research and routine set-ups. This is a powerful technique although subjective to the observer's experience for gating and evaluating the quality of data notably when performed as a routine test over long periods of time. Gating and analysis subjectivity can be overcome by high level of specific professional training and experience to produce consistent and reliable data and decrease the subjectivity-related variability that is observed in flow analysis data [677]

FC benefit from increasing range of commercially available reagents which has led to major improvement in staining protocols over the years [678]. However, this also has drawbacks. The use of different commercially available antibody clones, which often fail to identify the same epitope on a target protein [616], has been shown to contribute to data discrepancies for certain cell subsets between different labs/studies. This suggests critical consideration is needed for assessing data obtained from different settings.

Lack of standardization in assay and instrument setup has also been suggested as another area of concern. For instance, a lack of interlaboratory reproducibility exists due to variations in local experimental design and execution (e.g., choice of antibody clones, choice of fluorochrome-conjugated (usually using commercially available already conjugated),, sample preparation protocols, acquisition procedures) and instrument setup (e.g., cytometer performance and settings) which may arise at different levels of FC workflow [602, 678]. In clinical settings notably, the variability in methodological approaches to monitor cells of the immune systems in many disease can be used to illustrate this problem more specifically. This calls for a well-controlled FC standardisation underpinned by a

detailed set of SOPs peculiar to each disease, starting from the pre-analytical (sample preparation and staining protocol) to analytical (cytometer setup and acquisition) and then to post-analytical (gating and interpretations) phases [602, 678]. Improvements in the analytical workflow of the FC, as described in **Figure 8** will help to minimize data variability and optimize the quality of clinical flow cytometry given the current expanding role of the technology [602, 679].

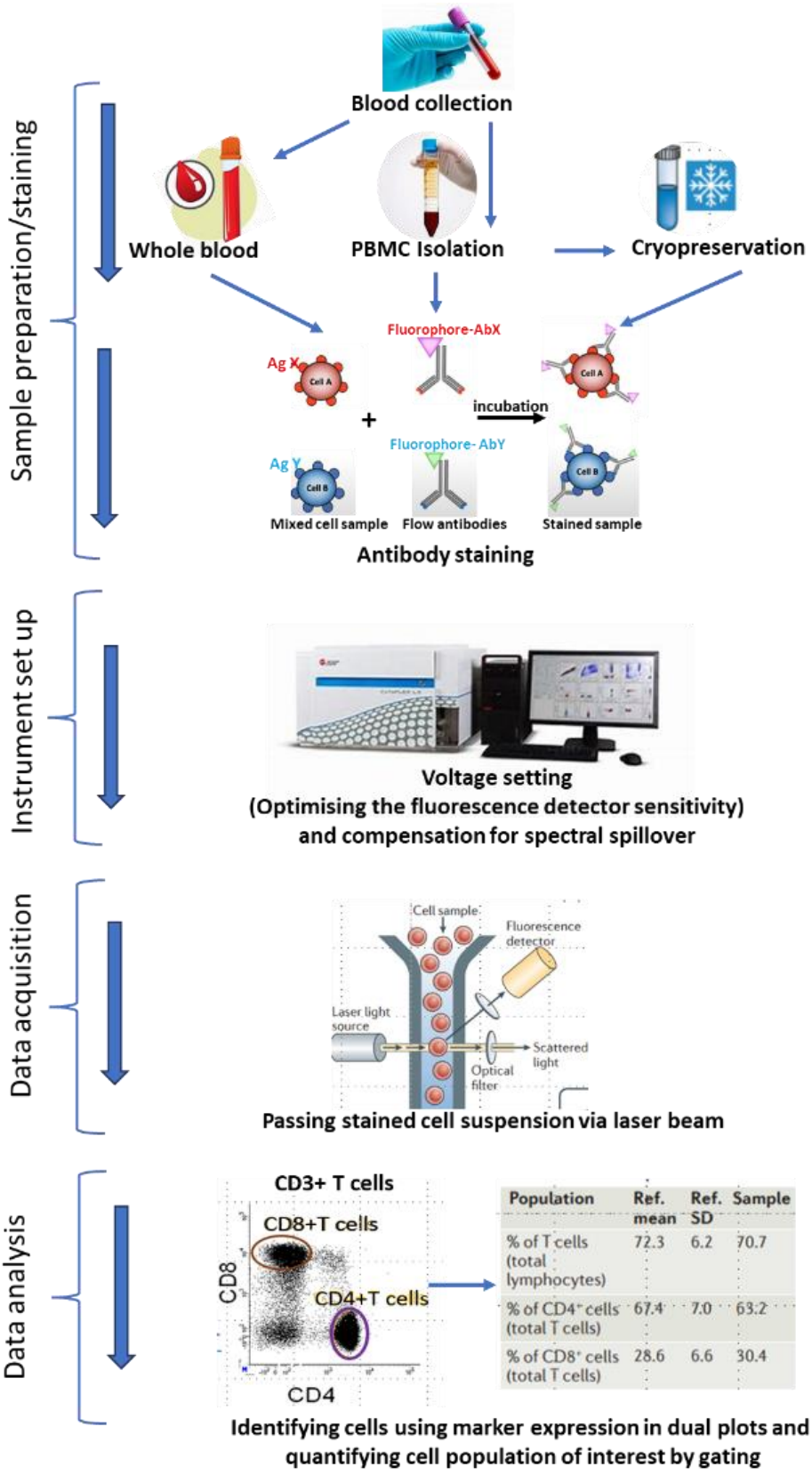

Sample preparation/staining
Blood collection
Whole blood
PBMC Isolation
Cryopreservation
Fluorophore-AbX
Ag X
Cell A
incubation
Fluorophore- AbY
Ag Y
Cell B
Mixed cell sample
Flow antibodies
Stained sample
Antibody staining
Instrument set up
Voltage setting
(Optimising the fluorescence detector sensitivity)
and compensation for spectral spillover
Data acquisition
Cell sample
Fluorescence detector
Laser light source
Scattered light
Optical filter
Passing stained cell suspension via laser beam
Data analysis
CD3+ T cells
CD8+T cells
CD4+T cells
CD8
CD4
Population
Ref. mean
Ref. SD
Sample
% of T cells (total lymphocytes)
72.3
6.2
70.7
% of CD4+ cells (total T cells)
67.4
7.0
63.2
% of CD8+ cells (total T cells)
28.6
6.6
30.4
Identifying cells using marker expression in dual plots and
quantifying cell population of interest by gating

Figure 8 A conventional flow cytometry workflow stages.

- **Sample preparation stage**- blood collection, separation of mononuclear cells, and sometimes cryopreservation, before staining, and sometimes whole blood staining, with fluorescent antibody conjugates.
- **Instrument setup stage-** setting voltage gains for the photomultiplier tubes (PMTs) for optimal sensitivity.
- **Data acquisition stage**- passing the stained cells via a probe in a single file through a laser beam and recording the fluorescence emission from all the bound antibody conjugates.
- **Data analysis stage**- cell populations of interest are defined by gating and then reported.

1.9.3 Successful use of FC in NHS

FC is nowadays available in NHS clinical immunology laboratories countrywide and used for several clinical measurements in the management/monitoring of different conditions . For instance, immunophenotyping analysis (identification and quantitation) of different cells and their subsets is one of the most important applications of FC which has an established role in diagnosis and disease monitoring of many diseases including haematology, oncology and inflammatory/ autoimmune diseases, etc [680, 681] (see **table 13**).

The technology has therefore been instrumental for translational research in many clinical situations [611] and has made the transition from a research tool to standard clinical testing.

Table 13 Flow cytometry in routine lab services

Flow cytometry test biomarker	Condition managed/monitored	NHS
CD4 count/Quantification CD4+ and CD8+ T cells	HIV- *human immunodeficiency virus* disease monitoring Haematology patients post-BM/PBSC transplant or on treatment	[682, 683]
Treg	Rule out IPEX (Immuno-dysregulation polyendocrinopathy enteropathy X-linked syndrome) (dysfunction of FOXP3)	[682, 683]
Naïve T-cells (CD3/4/45RA/27) as part of extended lymphocytes panel	PID (pelvic inflammatory disease) suspected patients	[682, 683]
CD3+T cell enumeration	Patients undergoing solid organ transplant and receiving OKT3 or ATG therapy	[682, 683]
B cell count	Rituximab monitoring- monitor patients who are on rituximab for the treatment of rheumatoid arthritis or certain haematological diseases (as directed by a haematologist) or	[682]
B-cell panel	PID, sometime post rituximab/HSCT (haematopoietic stem cell Transplant) patients	[682]
HLA-B27 positivity	diagnostic indicator for several autoimmune and immune-mediated diseases including ankylosing spondylitis, reactive arthritis and anterior uveitis.	[682]
Leucocyte Immunophenotyping	leukaemia and lymphoma diagnosis (Cell Markers), AML, ALL, MPD	[684-686]
Stem cell donor/enumeration	CD34+ cells	[686]
Flow crossmatch	Renal transplant	[687]
Neutrophil function	CGD- chronic granulomatous disease	
TBNK testing:(T-cells (CD3/4, CD3/8), B-cell (CD19), and NK-cells (CD16/56)	Suspected Primary/secondary immunodeficiencies or SCID (Severe combined immunodeficiency) and **Lymphoproliferative disorders**	[686]
measurement of CD55 andCD59 on red cell	GPI Linked Proteins for Paroxysmal Nocturnal Haemoglobinuria (PNH)	[683, 685, 686]
B-ALL CD38/CD123/CD58/CD22	Measurable residual disease (MRD) Flow Cytometry assessment	[686]

1.10 Dry tube based FC method

The conventional wet-tube (WT) based flow cytometry method currently in use in NHS settings (as well as most research labs) involves time-consuming multiple pipetting steps, which often results in having to exclude data due to mistakes (missing staining-antibodies and other technical errors). These are recognised limitations to using FC in routine that have been considered by industries like Beckman Coulter [688, 689] Cytognos SL [690] and Becton Dickinson (BD) [690-692] who developed solution to overcome this issues. BD has been an international leader in the FC field and developed a technology called dry tubes. This technological development was driven by the need to improve quality and productivity while reducing the costs/time of the workflow in health services flow cytometry laboratories. Dr Ponchel agreed to test some DT with BD on the backbone of the biomarker research program funded by LIRMM.

DT is a ready to use antibody-coated dry flow tube produced using the BD dry coating technology. The DT for the T-cell panel and Treg panel were designed for quantifying naïve-CD4+T, inflammatory related cells-CD4+T (IRC-CD4+T cell) and CD4+Treg which were the 3 most important blood cell subsets biomarkers used in RA patients at the time (2018-19). Two DT were produced for the validation study.

- One consists of a cocktail of 5- antibody (CD3, CD45RB, CD4, CD62L, CD45RA) dried in a 12 x 75mm tube for T-cell panel for CD4+T subsets
- The second was a 3- antibody cocktail (CD4, CD127, CD25) for Treg panel.

DT are intended to minimise lab procedures (i.e., handling mistake, missing Ab, etc…) and gain time when used in routine settings that would provide standardises processes notably for clinical trials as well as to disseminate a test for multicentre studies and facilitate adoption internationally.

The DT antibody clone selection and optimisation were performed between my lead supervisor, Dr Ponchel, and BD in 2019 . The project had reached the stage of performing a WET/DT data comparison on actual patient samples. Data were to be acquired as % of lineage with no need for mean fluorescent intensity (MFI) hence the fluorochromes were adapted by BD owned clones (commercial considerations) as well as based on technical issues with fluorochrome chemistry while antibody clones were chosen after comparing various clones for similar results during the optimisation phases.

The DT was then manufactured (n=50) using the same concentrations of antibodies that were used with the WT version during the comparison study. A figure of the CD4+T-cells profiling is attached (generated by my supervisor during the process, **see appendix A**).

Chapter 2

Hypotheses, Aims, and Objectives

Overall Aim

The significance of CD4+T subsets as biomarkers for the management of RA across the IAC has been established [1, 3, 5, 362, 520]. Data were obtained for other lymphocyte subsets, by the NHS immunology services but they have not so far been explored for their possible value in RA. Given that several transitions between the phases of the IAC are still lacking biomarkers, the general aim of my PhD is to investigate whether there is an added value in quantifying CD8+T-cell subsets, B-cell subsets, and NKT/NK-cell in addition to CD4+T-cells, with respect to association with disease progression outcomes across the IAC. As these panels were performed routinely by NHS services, the exploration of these flow cytometry panels may offer new biomarker tools to manage and stratify the risk of developing IA in ACPA+ at-risk individuals, predicting evolving to RA in EAC patients, and identify RA as well as PsA patients who may benefit from a specific cs- and/or b-DMARDs therapy.

The wet-base (conventional protocol using liquid antibody reagents) FC procedure classically used in research and NHS lab involves multiple antibodies pipetting which is time-consuming and prone to operational mistakes. Therefore, the second aim of my PhD was to develop a novel dry-tube flow-cytometry technology, with equivalent performances that would reduce the turn-around time compared to routine wet-tube flow-cytometry (as currently used by the NHS services). This will enable further the development of panels of clinical value for worldwide use.

My Overall hypothesis is that:

Lymphocyte dysregulation assessed by FC can provide biomarkers for the management of rheumatoid arthritis (RA). This is underlined by the sub hypotheses that assessing lymphocyte subsets using FC in hospital services is sufficiently reliable to have biomarker value.

Objective 1: Demonstrate the reliability of the FC data obtained by NHS services.

This would enable better strategies for the improvement in the quality of test results used for clinical decision-making on patients' management.

Objective 2: The Dry tube technology has the potential to replace wet-tube technology in daily routine flow cytometry (This is an industrial collaboration in an R&D project).

Specific aims and objectives

I. To validate the use of novel flow-cytometer dry-tube technology compared to wet tube currently used
II. Developing SOPs for the use of a new Kit by hospital flow-cytometry services

This would enable the adoption of pre-established panels for clinical use worldwide.

A second hypothesis is that:

Comprehensive blood cell phenotyping has clinical value for the management of RA. Predicting clinical outcomes across the IAC and providing further understanding of the events driving progression,

This is underlined by the sub-hypotheses that

I. quantifying CD8+T–cell subsets, NK-cell and NKT can provide clinical value.
II. B-cell quantification across the IAC can have predictive value.
III. combining these with the three CD4+T-cell subsets already known to be predictive can offer added clinical value.

Objective 3: To analyse 18 subset flow cytometry data with respect to clinical outcome in several groups of patients across the IAC.

This would enable improvement of the performances of current prediction models that are using only three CD4+T-cells subsets, in combination with demographic and clinical data and offer better prediction to manage and stratify the risk of progression across the IAC.

A third hypothesis was that:

The analysis of frozen samples in a clinical trial of Psoriatic arthritis (PsA) can provide useful biomarkers value.

This would offer new tools to discriminate between patients who will benefit from such treatment or not, leading to personalised and cost-effective treatment in PsA.

Due to the impact of COVI-19, this aspect of my work could not be completed as planned (see covid-19 statement **section 6.1 page 355).** I could not get access to the clinical data to merge with the flow data to achieve this objective. Therefore, my supervisors and I designed an alternative objective for me to exploit the flow data acquired. This led to,

Objective 4: Flow cytometry can bring valuable information for retrospective studies using frozen samples.

specific aims and objectives

I. To acquire and analyse flow-cytometry data from frozen samples.
II. To understand the limitation of working with frozen samples
III. To develop SOPs for studies where samples have to be frozen.

This would provide validation of the FC biomarkers using a frozen sample. Thus, would enable a better understanding of the impact of freezing on immunophenotyping and the development of SOP in the context of inflammatory arthritis disease.

In pursuance of robust flow biomarkers project in autoimmune inflammatory arthritis notably, RA, my PhD work seeks to phenotype additional novel blood cell biomarker(s) across the IAC, notably with respect to association with disease progression from one stage to the next as well as response to treatment using flow-cytometry as the main technology.

Chapter 3

Materials and Methods

3.1 Ethical approval

As a part of the early Inflammatory Arthritis (IA) programme of work in Leeds Institute of Rheumatic and Musculoskeletal Medicine (LIRMM) groups, patients were recruited from the wider regional primary care services into several studies, one being the cyclic citrullinated peptide (CCP) study and the second, the early arthritis clinics (EAC). Blood samples at inclusion into the studies were sent to the NHS-immunology services. Cell subsets data were acquired by NHS services.

Patients' data utilised in this project were therefore retrieved from those 2 longitudinal studies:

- The Leeds Inflammatory Arthritis Disease Continuum the "(IACON register", REC approval: 09/H1307/98)
- The Coordinated Programme to prevent Arthritis study (the "At-Risk CCP register", REC approval: 06/Q1205/169) both comprising >5000 participants enrolled between 2008-2022.

Other patients' data used were from ongoing clinical trial:

- An investigator-initiated double-blind, parallel-group randomised controlled trial of GOLimumab and Methotrexate versus Methotrexate in very early Psoriatic Arthritis (PsA) using clinical and whole-body MRI outcomes: the GOLMePsA study (REC approval:14/EM/0124).

Ethical approval for data recording and sample storage was obtained from Leeds Teaching Hospitals NHS Trust Ethics Committee for the study (see approval letters, **appendices B-D**), and all participants provided informed written consent.

The raw data (FCS files) for patients across the IAC was made available for use in the current study whereas frozen sample of patents from GOLMePsA study were used in the current study. Flow cytometry panel design for data acquisition for these samples were designed by my lead supervisor (FP) before my PhD while data were acquired by me.

My thesis is composed of a number of patients cohorts selected based on availability of clinical data needed for the specific study aim.

3.2 Patient group description

3.2.1 The Inflammatory Arthritis Continuum (IAC) study

The study across the IAC included different clinical groups of patients with various outcomes of interest to my study, comprising the following cohorts summarised in **Table 14;**

3.2.1.1 The at-risk cohort

In the at-risk cohort, lymphocyte subset phenotyping at baseline was investigated for the prediction of progression to Inflammatory Arthritis in anti-citrullinated protein antibodies (ACPA) positive at-risk individuals (the at-risk cohort). A total of over 1500 data for at–risk individuals were available. However, following data cleaning and patients' selection, current study cohort comprised progressors (n=93) versus non-progressors (n=117) and various control (healthy and family members, n=180) were included to established reference range. Patients were recruited from regional primary care services as previously described [1, 3, 312] based on positivity for a laboratory test for the presence of anti-citrullinated protein antibody (ACPA) or rheumatoid factor (RF). Briefly, patients were included based on a new musculoskeletal complain (including rotator cuff tendonitis,

subacromial bursitis, carpal tunnel syndrome, tendonitis, back pain, or epicondylitis) and ACPA positivity (using various CCP-2 tests over the years, Axis Shield Diagnostics Ltd, Dundee, Scotland and the BioPlex 2200 kit, BioRad, USA, performed by NHS services) or a RF performed using NHS standards protocol. If positive, they were referred to secondary care for clinical evaluation. If evidence of clinical synovitis were detected patient were transferred to our EAC. In case of absence of clinical synovitis, patients were enrolled into the At-risk CCP register. Progression was defined by the development of clinical synovitis evaluated by a senior rheumatologist comprising at least 1 swollen joint. Follow-up has been ongoing since 2008.

3.2.1.2 Early arthritis clinic (EAC) cohort

In the early inflammatory arthritis cohort, lymphocyte subsets phenotyping at the first visit were investigated for the prediction of progression to RA in patients with early symptoms of clinical synovitis.

This cohort (n=306) included patients with early symptoms of inflammation in joints, who were classified for RA, using the EULAR 2010 criteria (n=206) versus another form of arthritis (non-RA, n=100) over 2 years of follow-up since referral. These were disease-modifying anti-rheumatic drugs (DMARDs)-naïve patients, but the use of non-steroid analgesia and intra-muscular injection of steroid (depometrasone) within 3 months of the time of referral to the EAC, were permitted.

3.2.1.3 Methotrexate (MXT)-treated early RA cohort

Here, lymphocyte subsets phenotyping were analysed for the prediction of MTX-induced remission in RA patients treated with a Treat-to-target approach, starting with MXT as first line drug a recommended by National Institute for Health and

Care Excellence (NICE) guidelines. In this cohort, patients (n=205) were treated with MTX dose initially using a Treat-to-target approach [332] then escalated to high doses or addition of other synthetics DMARDs for instance Sulfasalazine (SSF) or Hydroxychloroquine (HCQ) if not responding to MXT alone. Patients achieving MTX-induced remission (n=106) versus non-remission (n=99) based on DAS28<2.6 to define remission at 6-months, were selected on basis of available flow cytometry data (notably over COVID-19 pandemic).

3.2.1.4 Remission First Visit cohort

Early RA patients treated with MTX and other synthetics that achieve remission for the 2 consecutive visits over 2 years of the first visit in IACON. Blood collected at for the first visit in remission (independently of how long it took to get there) was used to get the flow data.

These clinical groups of patients are representative of similar cohorts at each stage of the IAC

3.2.2 Frozen samples from GOLMePsA trial

The GOLMePsA study is an investigator-initiated double-blind, parallel-group randomised controlled trial of GOLimumab + Methotrexate versus Methotrexate. The cohort included a total of 84 PsA patients classified according to CASPAR criteria [693], with up to 24-month symptom duration with at least 3 swollen and 3 tender joints or 2 swollen and 2 tender joints and one tender enthesis, recruited during routine appointment at our EAC. These were treatment naïve-PsA patients randomised on a 1:1 basis treated with a combination therapy (Golimumab (GoL) +Methotrexate (MTX)) versus MTX monotherapy+ placebo for a total of 24 weeks duration (primary endpoint at which Golimumab was discontinued in patients with combination therapy) while MXT was continued in both arms until week 52, the

final endpoint of the study. Patients who consented to the biomarker sub-study provided blood samples at baseline and weeks 12, 24, 36 and 52 and PBMC separated by the tissue bank technician at Chapel Allerton Hospital using SOP (**See appendix E**). These PBMCs were stored at -150^{0}C and later made available for the current flow cytometry work included in the protocol.

3.2.3 Healthy Controls (HC)

Healthy controls (HC) (n=180) were recruited mainly as volunteers from laboratory staff. These were used to establish reference ranges for cell subsets data and update age relationships where relevant as previously described [4] .

Table 14: Summary of data source and study cohorts used in this project

IAC study cohorts	Leeds observational/trial cohorts	Total (n)	Outcome studied (n)	Study start	Study end
Pre-clinical	CCP	~1500 (selected 210)	Progressors n= 93 No Prog n= 117	2008	ongoing
EAC	IACON RADAR	~2500 (selected 306)	RA n= 206 Non-RA n= 100	2010 2015	2015 ongoing
MTX study	IACON RADAR	~ 500 (Selected 181 + 24 from below= 205)	Remission n= 106 Not in rem n=99		
Remission	IACON-rem	~144 (Selected 24)	Flare No flare (Not part of this study	2008	ongoing
PsA samples	GOLMe PsA	Planned for 84	Good response Poor response (No access to data)	2018 (stopped)	Last patient / last visit March 2024
HC samples	variable	180	Not applicable	2008	ongoing

3.2.4 Validation of novel Dry-tube assay as a new flow cytometry method

For this study, fresh blood samples from active RA patients (n=44) were utilised for the technical validation of a flow-cytometer kit based on dry-tube technology. Here, dry tube CD4+T cell subsets (naive/Treg) phenotyping was performed and compared with data from wet lab retrieved from NHS-immunology lab for the same patients.

3.2.5 Flow cytometry validation on frozen samples

The aim of this analysis was to ensure the appropriateness of frozen PBMC for LS qualification. Here, lymphocytes subsets were phenotyped using frozen samples from early, drug-naïve with inflammatory arthritis included in the GOLMePsA clinical trial (see above).

3.3 Flow cytometry data retrieval from NHS service

3.3.1 Flow cytometry data files searches

FCS file of existing raw flow cytometry data from panels that were developed to predict the progression alongside the Inflammatory Arthritis Continuum (IAC), were retrieved from NHS servers and utilised in the current study. The flow cytometry was performed using routine standard techniques followed by the NHS-Immunology flow services (SOP, **appendices F-I)**. The raw data was made available. I had no participation in this wet laboratory data acquisition phase, for data derived from patients in the IAC.

Given that the study involves human subjects, the obligation to respect and protect the right of participants was given particular attention. Data (flow cytometry files) were transferred securely unto password specific NHS-encrypted external hard drives and University password protected N-drive which are only accessible

to the research team. To maintain patient's confidentiality, all files' coding were anonymised by removing patients' names (which is the normal practice in NHS file naming), and study codes were used instead to enable data allocation to the right study group. This did not threaten the participant's confidentiality as data were presented and reported as aggregated data findings from the overall group/cohort. Some of the files were however annotated with patients' names rather than study code/visit time point (when a study code was not clearly provided). This complicated my work as I have no direct access to personal information, all register data being anonymised. Data were sorted through a phase of "detective work", whereby clinical colleagues were able to allocate a patient's study code so that the FCS-file could be renamed.

Each FCS file was then double-checked for time-point either baseline week zero (BL/W0) or follow-up visits, 3, 6, 9 or 12 months against the date of the visit to clinic to match the data of that visit in the clinical records to ensure that correct FCS files were included in the right cohort for analyses.

3.3.2 Clinical data retrieval

Clinical data collected as part of the register (at-risk CCP+ or EAC study) were downloaded from the several NHS-databases (IACON/CCP+ register), included demographic (age, gender, duration of symptoms, smoking habit), inflammatory markers (C-reactive protein (CRP), erythrocyte sedimentation rate (ESR)), genetic (presence of Shared epitope (SE)), autoantibodies status (anti-citrullinated protein antibodies (ACPA), and rheumatoid factor(RF)), and physical assessment (Joint tenderness, Joint pain, and early morning stiffness (EMS)). 28-joint disease activity score (DAS28) for patients were recalculated for individual parameters using the DAS (4 components) CRP formula:

DAS28-4(CRP) =0.56*SQRT(TJC28) +0.28SQRT(SJC28) +0.36*ln (CRP+1) +0.014*GH+0.96).

The 78-joint disease activity score was used in the at-risk CCP+ register.

3.4 Flow cytometry analysis

Raw flow cytometry data (FSC file) from 5 panels were retrieved as electronic files. Flow cytometry data acquisition using fresh sample was performed on a FACSCanto flow cytometer (Beckon Dikson, United State Amercian) by NHS, and data analysed using BD Biosciences FACS-DIVA software version 8.0 by me.

Antibody clones and staining protocols used are described in detail in NHS immunology lab SOP (**appendix F-I)** The gating strategies are described in **Figure 9**. Blood subsets frequencies were reported as percentage of the parental population.

Due to subjectivity of technique, training and validation of flow cytometry analyses skill were carried out, and preliminary data (over 50 data points for each panels) compared to that of my supervisor with over 20 years of flow cytometry experience. Until satisfactory comparable data were obtained (rho > 0.950 by spearman correlation). Additionally, 30 data points for T-cell subsets and Treg panel from NHS–lab also showed comparable data (rho > 0.933) with the independent analysis performed.

Next, eighteen (18) subsets were acquired and quantified by me using frozen sample from these 5 panels as listed in **Table 15**

Table 15 : 18 lymphocyte subsets analysed

Panel	Markers used	Subset quantified	Age-related
1. Lineage	CD19+, CD3+, CD4+, CD3+, CD8+, CD3-CD56high/low, CD3+CD56+	B CD4, CD8, NK $^{bright/dull}$ NKT	No
2. CD4+T cells	CD3+, CD4+, CD45RA+, CD62L+ CD3+, CD4+, CD45RA-, CD62L- CD3+, CD4+, CD45RA+, CD62L-	Naïve, memory IRC	Yes, (Naive/ Memory)
3. CD8+T cells	CD3+, CD8+, CD45RA+, CD62L+ CD3+, CD8+, CD45RA-, CD62L- CD3+, CD8+, CD45RA+, CD62L-	Naïve, memory IRC	
4. B-cells	CD19+, CD27-, CD38- CD19+, CD27+,CD38- CD19+CD24+, CD38+high CD19+CD27+, CD38+	Naïve, Memory, Putative Breg plasmablasts	Yes, (naïve/ memory)
5. Th17 cells	CD3+, CD4+, CD161+, CCR6+, CXCR3-,	Not applicable	No

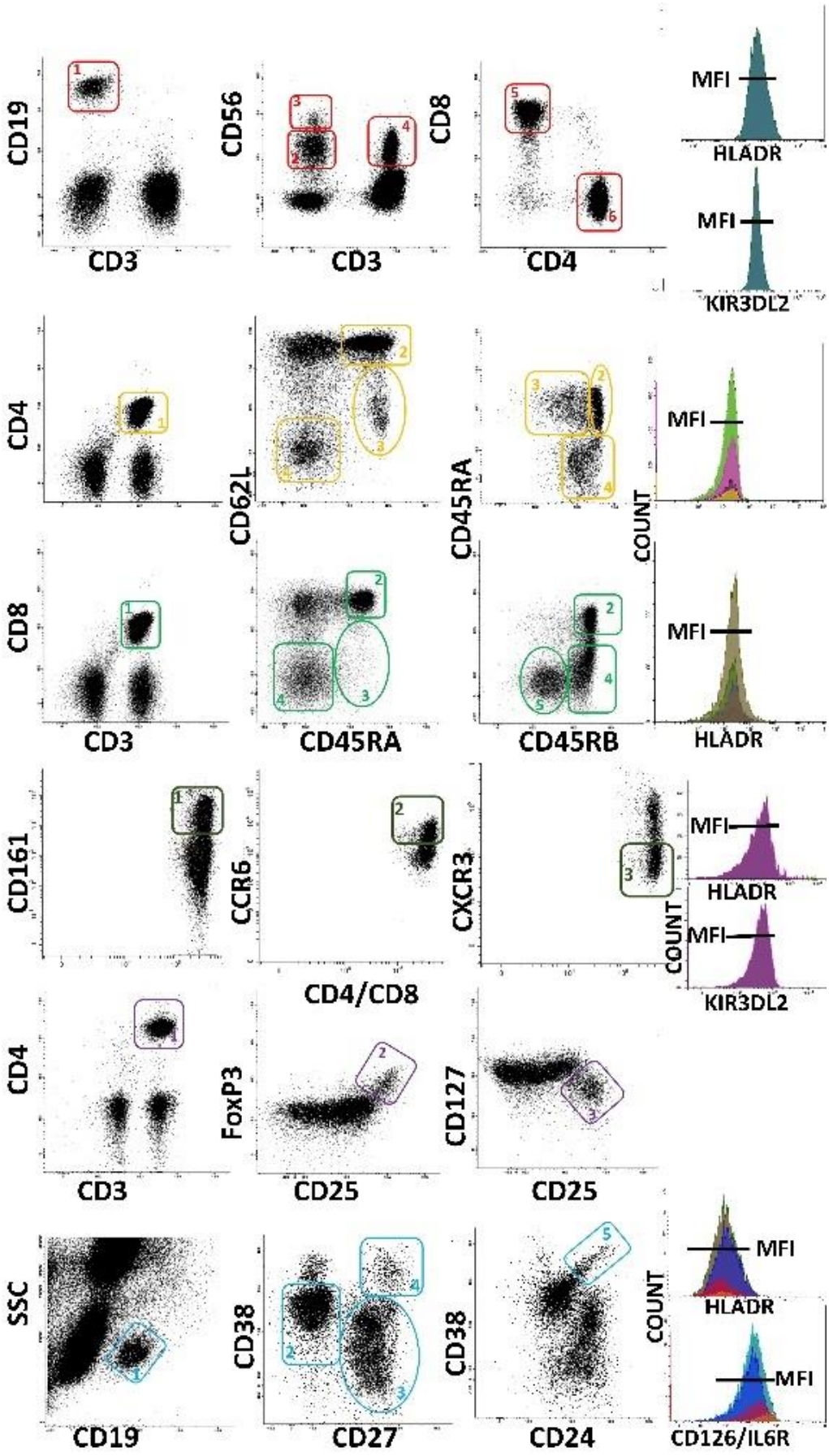

CD19
CD3
CD56
CD3
CD8
CD4
MFI
HLADR
MFI
KIR3DL2
CD4
CD62L
CD45RA
MFI
COUNT
CD8
CD3
CD45RA
CD45RB
MFI
HLADR
CD161
CCR6
CD4/CD8
CXCR3
MFI
HLADR
COUNT
MFI
KIR3DL2
CD4
CD3
FoxP3
CD25
CD127
CD25
SSC
CD19
CD38
CD27
CD38
CD24
COUNT
MFI
HLADR
MFI
CD126/IL6R

Figure 9 Gating strategies for lymphocytes cell and subset.

Lymphocyte lineages were quantified for B-cells, NK cells subsets, NKT-cells CD4+ and CD8+T-cells. Lymphocytes populations were first identified through a FSC-A vs SSC-A dual plot. The lineage subsets included red gates: gate1- B-cells were identified based on CD19 expression, gate 2- NK cell $CD56^{dim}$, and gate- 3 NK cell $CD56^{bright}$ were identified based on low and high expression of CD56 on a dual plot for the expression of CD3 and CD56 while gate 4- NKT cell were defined as double CD3+CD56+. The CD3+T-cells were further interrogated on a dual plot for gate 5- CD8+T cell and gate 6- CD4+T-cells. Result was expressed as percentage (%) of total lymphocytes, the sum of each lymphocyte cell counts (as 100% events) being used to exclude possible contamination of monocytes in the original lymphocyte gate. The HLADR/ KIR3DL2 expression on the 6 lineage subsets was analysed recording the MFI.

For CD4+T-cells subsets green gate: gate 1-CD4+T were analysed for subsets, gate-2 naïve cells, gate 3- inflammation related cells (IRCs), and gate 4- memory cells based on the expression of CD45RA and CD62L. CD45RB was used to refine the naïve (gate-2) and IRC (gate-3). Results for each individual subset were expressed as percentage (%) of total CD8+T-cells. The HLADR expression on the 6 subsets was analysed recording the MFI.

For CD8+T-cells green gate : gate 1-CD8+T were analysed for subsets, gate-2 naïve cells, gate 3- inflammation related cells (IRCs), and gate 4- memory cells based on the expression of CD45RA and CD62L. CD45RB was used to refine an additional population of expended-CD8 cells (gate-5). Results for each individual subset were expressed as percentage (%) of total CD8+T-cells. The HLADR expression on the 6 subsets was analysed recording the MFI.

For CD4+Th17 cells grey gate: Th17 cells was identified based on CD161+CCR6+CXCR3-, a 3-marker identification system. CD4+ T-cells gated from a lymphocyte forward/scatter size were analysed first for gate-1, the expression of CD161. CD161+ cells then were analysed for gate 2, for the expression of the second marker CCR6. Finally, gate 3, th17 cells were defined based on negative expression of CXCR3 on CCR6+CD161+ cells. Result was recorded as % of CD4+T-cells. The expression of HLADR/KIR3DL2 on Th17-cells was then analysed recording MFI.

For CD4+Treg-cells purple gate: gate 1-CD4+T were analysed for gate-2, regulatory T-cells (Treg) based on the high expression of CD25 and FoxP3. CD127 was used to refine this Treg (gate-3) based on low expression of CD127. Results were expressed as percentage (%) of total CD4+T-cells

For B-cell subset phenotype blue gate. Gate-1, B-cells were first gated based on the expression of CD19 in a dual plot with the scatter parameter. Gate-2, naïve B-cell, gate-3, memory B-cell, and gate-4 plasmablasts (PBs) were identified based on the expression of CD27 (positive or negative) and CD38 (negative, positive and high). Gate-5, putative Breg were identified as included in the transitional-2 B-cell population using high expression of CD38 and positivity for CD24. The expression of HLADR/ILR6 on Th17-cells was then analysed recording MFI. Results for each individual subset were expressed as percentage (%) of total B-cells.

3.5 Normalisation of data for subsets dependant on age

Age-relationship for specific cell subsets have previously been described in health over 6 decades of age using healthy controls [4]. Updating previous data with additional data (n=60) to refine the age normalisation (**Figure 10**). Age range was done over time(n=120) [5], and I used (n=180) for all analysis performed in my thesis. The addition of more healthy controls increased the significant of age normalisation to reach significant rho values (**Table 16**). This affected 7 subsets as previously described (naïve and memory CD4+T- cells, regulatory CD4+T-cells (Treg), naïve and memory CD8+T-cells, and naïve and memory B-cell). Other subsets were not related to age.

This approach was chosen to ensure that meaningful comparisons on subsets between outcome groups across IAC are attributable only to the effect of the disease. This allowed for ranges of data to be expected in health to be directly compared to disease and to define 95% confidence limit (Cl) of distribution. As show in **figure 11**, data are not fitting normal distribution for all the 18 subsets.

Table 16 Correlation Analysis between Age and lymphocyte subsets

Subsets	Spearman correlation coefficient, rho	Coefficient of determination, R^2	Correlation equation [X= age]	p-value
Naive CD4	-0.717	0.516	Y = -0.5461*X + 63.78	<0.0001
Memory CD4	0.323	0.131	Y = 0.1538*X + 10.05	0.0003
Treg	0.547	0.321	Y = 0.0605*X + 1.861	<0.0001
Naïve CD8	-0.703	0.559	Y = -0.7635*X + 56.76	<0.0001
Memory CD8	0.521	0.279	Y = 0.4816*X + 16.04	<0.0001
Naïve B	0.660	0.422	Y = 0.6254*X + 38.35	<0.0001
Memory B	-0.618	0.372	Y = -0.5884*X + 55.54	<0.0001

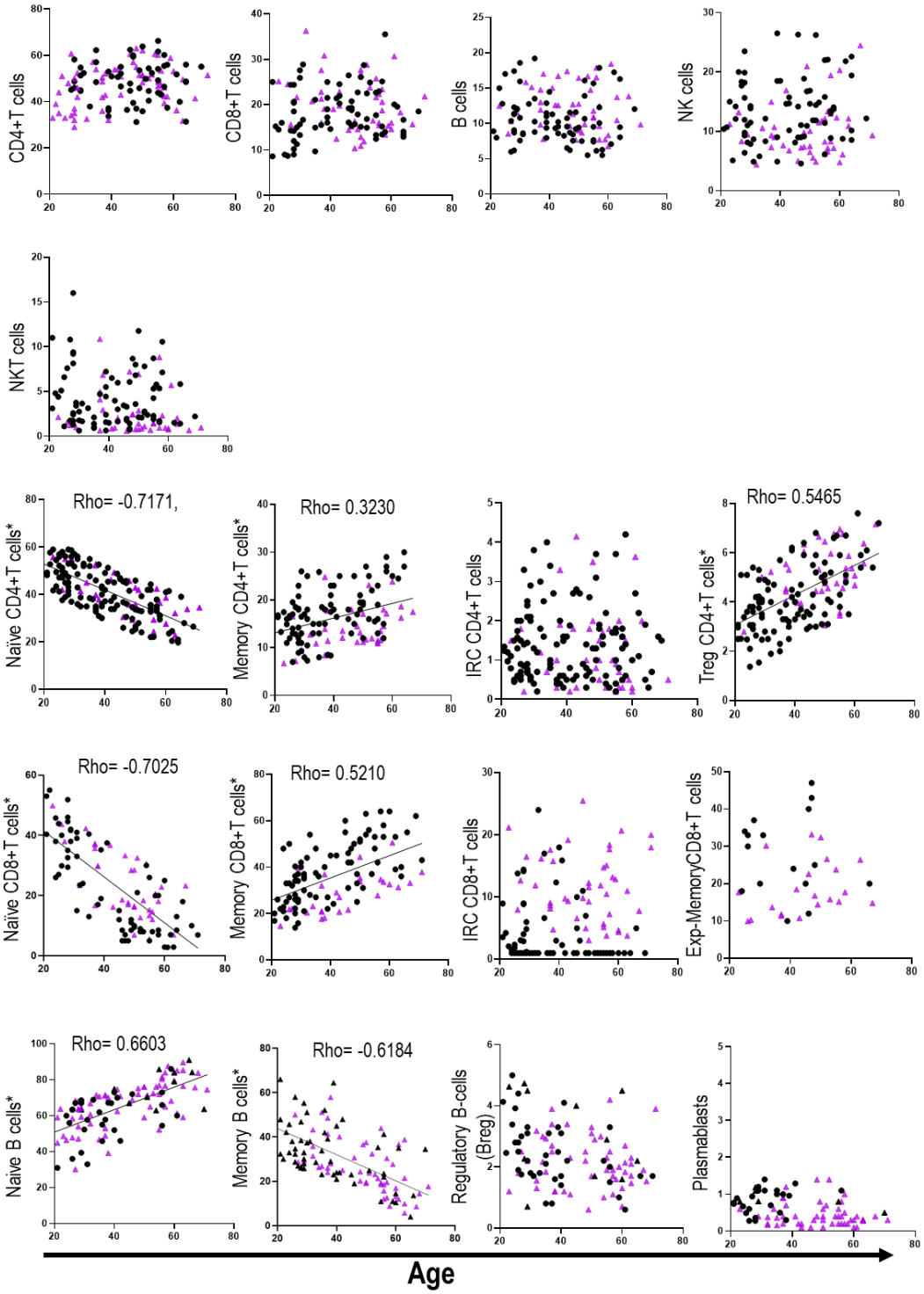


Figure 10 Association of 18 lymphocyte subsets in relation with age.

Cell subsets were quantified as percentage (%) of total parental lineage (CD4+T-cells, CD8+T-cells and B-cells) and data presented in correlation to age of the HC (n= 180). Data obtained were in 2015 (black triangle) in addition to new data acquired since (purple triangle). Spearman correlation analysis performed confirmed significant age associations with 7 subsets (naïve and memory CD4+T- cells, regulatory CD4+T-cells (Treg), naïve and memory CD8+T-cells, and naïve and memory B-cell).

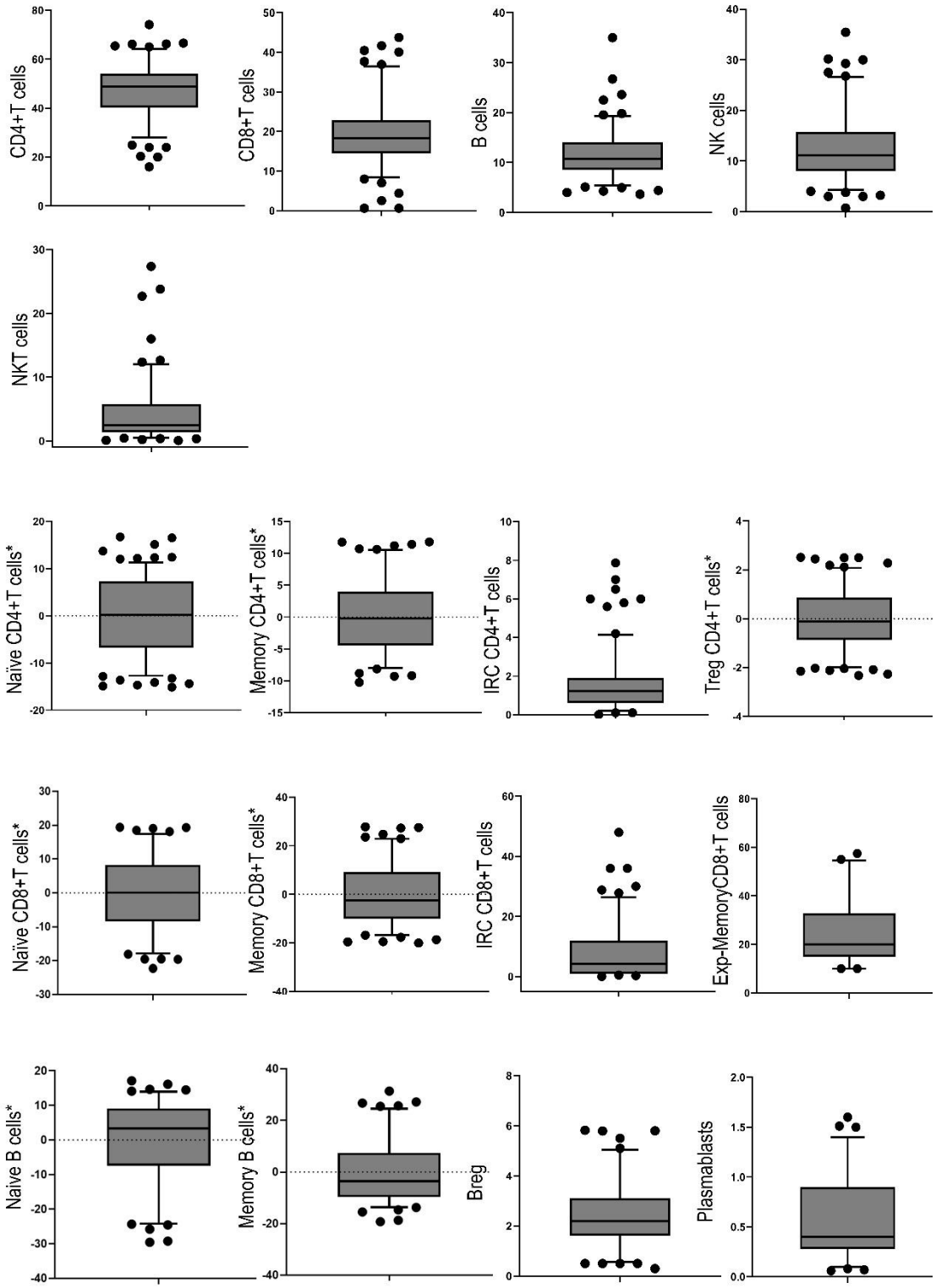


Figure 11 Data distribution in healthy control (HC).

Data acquired in 180 HC are displayed (after age normalisation, for 7 subsets (naïve and memory CD4+T- cells, regulatory CD4+T-cells (Treg), naïve and memory CD8+T-cells, and naïve and memory B-cell) as boxplot. Distributions are close to normal but not fitting. Hence data were treated as not normally distributed. Box plot displayed the median as the middle line, the interquartile range as box, and 95% confidence interval (CI) extremes as whiskers. Dots outside of whiskers represent the few points out of the 95%CI.

3.6 Flow cytometry data acquisition

3.6.1 Cells

Two different types of cells were used in the different panels of the work

- Frozen PBMC was used in wet tubes only
- Fresh blood (whole) was used in wet tube and dry tube experiment

3.6.2 Panel optimisation and compensation on frozen samples

3.6.2.1 Frozen cell panels development

Four-panels using up to 8 fluorochromes in various antibody combination were designed to quantify frequencies of several subsets (**Table 17**).

The antibodies chosen for the panel design were based on first, defining subsets of interest in this study and their associated markers to be detected **(Table 15**) and then matching right antibody-fluorochrome conjugate with the right markers to be identified on cell population of interest. The matching procedure, however, was based on expression level of markers (the antigen density), where makers with low expression (dim markers) were assigned to bright fluorochrome and vice visa which ensured better resolution of cell population. The fluorochrome spillover was minimised by using the "fluorochrome resolution ranking" and "spectrum viewer" which helped me assess the cross-laser excitation and fluorochrome spill over information [694, 695]

For the development of each panel, 10 flow tubes were prepared each containing 50ul of the Peripheral blood mononuclear cell (PBMC) (approximately 1-2 million cells). Cells in the first 8 tubes were stained with a single antibody using Ab volume as described in **table 17**. The 9^{th} tube was stained with all the antibodies of the panel while the 10^{th} tube was the unstained as negative stain control.

After 30 min incubation at 4^0C in the dark, cells were washed once with 1mL of phosphate-buffered saline (PBS) solution (for detail of preparation see Appendix M). After centrifugation, the supernatant was decanted, and pellets re-suspended in 300uL FACS buffer and kept on ice in the dark.

Samples were acquired using a 13 Detectors, 4 Lasers CytoFLEX S VA-B4-R3-12 Flow cytometer (Beckman counter, USA). Instrument setup in addition to panel compensation was done manually.

Each panel was adjusted for cell size and granularity using the unstained control tube setting voltage for the no stain lymphocyte population below the third decade of the fluorescence log scale. Fluorochrome compensation was established using single colour control tubes individually. Last all the antibody tube was used to define the gating strategy identifying the subsets to be quantified.

3.6.2.2 Frozen cell panels cell surface staining

Blood sample from PsA patients in the GOLMePsA clinical trial were processed for PBMC separation and cryopreserved in 1mL cryo-vials, stored in -180^0C were made available to me for the study.

Frozen PBMCs were thawed quickly at 37^0C, until a frozen solid agglomerate and can be transferred from the vial by tipping it into 9mls of PBS containing 2ul/1ml DNase solution [1 vail of DNase (100mg) dissolved in 2.75ml of PBS] and mixed by inverting for 3 times followed by centrifugation (5 minutes, 500g at 12^0C) (**appendix J**). The use of DNase was to help to eliminate free DNA from dead/damage cells that leads to cell clumping or aggregation. The supernatant was decanted, and the pellet re-suspend in 1mL PBS/DNAse solution and spun at 500g, 5 mins at RT. The supernatant was decanted, followed by drying of pellet by inverting the tube on towel for few seconds. Cells were re-suspended in

30uL blocking buffer containing DNAse (at concentration of 2ul of DNAse per ml of blocking buffer).

Aliquot of 50ul of cells were transferred into flow cytometry each tube before adding antibodies according to volume described in **table 17**. To insure consistence in pipetting, antibody cocktails were prepared weekly for each panel and stored at 4^0C. Incubation was done at 4^0C for 30mins in the dark. Excess unbound antibodies were washed off by adding 3mL of PBS/DNAse solution and centrifuged as above. The supernatant was decanted and cells re-suspend in 300uL FACS buffer containing 2µl/ml of DNase for analysis and kept on ice. Data were acquired using a Cytoflex S flow cytometer (VA-B4-R3-12 Flow cytometer, Beckman counter, USA).

Flow cytometry analysis was performed for the quantification of the LS subsets, CD4+T cell subsets, CD8+T cells, B-cell subsets, and CD4+Th17 cells using CytExpert software. Gating strategies have been previously described above (**Figure 9, page 126**)

The dual plot of FSC-width versus FSC-height which was also double-gated in a plot of SSC versus FSC was used to separate viable from dead cells (**Figure 12**).

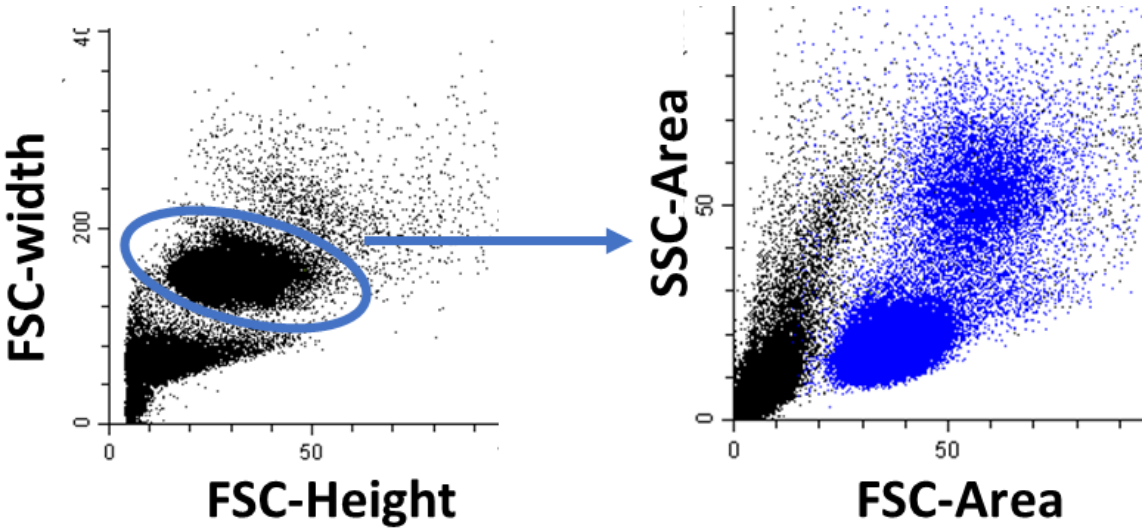


Figure 12 Gating strategy for identification of viable cells.
A representation of a dot plot showing the surrogate (FSC-width versus FSC-height dual plot) approach for viable cell identification. The viable cells (live cells) population (bule gate) was defined using the dual plot of FSC-width versus FSC- height. As dead cells and debris have reduced light scattering, Viable cells selected (blue gate) also double-gated in a plot of SSC Versus FSC to separate viable (blue region) from dead cells (dark region) .

Table 17 Panel showing Antibody clones / fluorochrome (clone name and company) used in each panel for frozen PBMC analysis

Fluoro chromes	Lymphocytes panel	µl	T-cells panel	µl	Th17 panel	µl	B-cells panel	µl
V421 (Clone) (Company)	CD4 (RPA-T4) BD	3	CD4 (RPA-T4) BD	5	CD4 (RPA-T4) BD	5	CD19 (HIB19) BD	10
V500 (Clone) (Company)	HLA-DR (G46-6) BD	5	HLA-DR BD	5	HLA-DR BD	5	HLA-DR BD	5
FITC (Clone) (Company)	CD3 (UCHT1) BD	7	CD3 (UCHT1) BD	7	CD3 (UCHT1) BD	7	CD24 (ML5) BD	15
PE (Clone) (Company)	KIR3DL2 (539304) R&D	10	CD45RA (HI100) BD	10	KIR3DL2 (539304) R&D	10	CD27 (M-T271) BD	10
APC (Clone) (Company)	CD56 (NCAM16.2) BD	5	CD126 (M5) BD	5	CD161 (DX12) BD	15	CD126 (M5) BD	5
PerCP Cy5.5 (Clone) (Company)	CD14 (MPHIP9 (also known as MφP-9)) BD	5			CCR6 (11A9) BD	5	IgM (G20-127)	5
A700 (Clone) (Company)	CD8 (RPA-T8) BD	5	CD8 (RPA-T8) BD	5	CD8 (RPA-T8) BD	5	CD38 (HIT2) BD	10
PE Cy7 (Clone) (Company)	CD19 (SJ25C1) BD	5	CD62L (DREG-56) BD	5	CXCR3 (1C6/CXCR3) BD	5	IgD (IA6-2 (also known as δIA6-2)	5
Cocktail volume (µl)	Mixed Abs	45	Mixed Abs	42	Mixed Abs	57	Mixed Abs	65

V421 = Brilliant Violet 421, V500=Brilliant Violet 500; FITC= Fluorescein Isothiocyanate; PE= phycoerythrin; APC= Allophycocyanin; PerCP Cy5.5= Peridinin chlorophyll protein-Cyanine5.5; A700= Alexa Fluor 700; PE Cy7= phycoerythrin Cyanine7; µl = microlitre, Abs= antibodies

3.6.3 Panel optimisation and compensation on fresh blood sample

3.6.3.1 Fresh cell panels : Development of Dry Tube experiment

Two-panels using up to 7 fluorochromes in various antibody combinations (5-colours 5-antibodies tube for T-cell panel and 3-colours 3-antibodies tube for Treg panel) were designed for the Dry Tube technology in collaboration with BD to allow to quantify frequencies of subsets. Not all the clones used in wet flow were from our collaborator BD. As such clones were chosen as replacement (before the start of my PhD) and the antibodies (Abs) used compared to Abs used in wet tube are described in **table 18.**

For the development of each flow panel, fresh whole blood (6 mL) from health individuals was collected into EDTA vacutainers tubes and inverted 8–10 times. Samples were collected in the morning and processed within 4 hours, as recommended [696]. Three millilitre (3mL) whole blood was added to 42mL of a of red cell lysis (RCL) buffer solution (see **appendix K**). The blood + RCL buffer mixture (opaque initially) was invert gently (about 5-10 mins) until it became more transparent, then centrifuged at 500g for 7 mins at 18^0C. The supernatant was decanted, and residual fluid removed by inverting the tube on absorbing paper. The pellet was re-suspended in 45ml RCL buffer, left for 1-2 mins, inverted few times and then spun as above. After decanting the supernatant, the pellet was re-suspended in 45ml PBS and centrifuged as before. The pellet was re-suspended in blocking buffer (see **appendix K**).

Nine (9) flow tubes were prepared each containing an aliquot of 50uL of the re-suspended fresh cells (at the end of the previous steps) (approximately 1-2 million cells.

Due to the effect of cell density on staining, manufacturers usually recommend antibody dilutions based on a fixed number of cells, approximately one million cells per sample. This guideline is especially crucial for staining low-density markers like CD127, where using a high number of cells can significantly decrease signal. This loss is due to the antibody being too dilute to bind all of the available binding sites [697]. Cells in the first 7 tubes were stained according to manufacturer's guidelines with a single antibody using Ab volume for the Wet tube as described in **table 18**. The 8^{th} tube was stained with all the antibodies of the panel while the 9^{th} tube was unstained as negative stain control. After 30 min incubation at 4^0C in the dark, cells were washed once with 1mL of phosphate-buffered saline (PBS) solution After centrifugation, the supernatant was decanted, and pellets were re-suspended in 300uL FACS buffer and kept on ice in the dark.

Samples were analysed using a 21 Detectors, 4 Lasers CytoFLEX VA-B4-R3-12 Flow cytometer (Beckman counter, USA). Instrument setup in addition to panel compensation was done manually as described above.

3.6.3.2 Fresh cell flow panel: Wet tube cell surface and intracellular staining

Six (6 mL) of fresh blood was collected as described above from active RA patients and transported using NHS service to WTBB lab at St James Hospital from the Early Arthritis Clinic in Chapel Allerton Hospital. Blood was collected between 8am to 12pm, packaged and transported through the blood shuttle service to St James Hospital the same day (kept at room temperature). Most sample were delivered to NHS immunology laboratory at least 3-4hrs after collection hence not before late afternoon. The wet tube staining was performed by the NHS-Immunology lab the next morning. Left over blood was forwarded to

me for a Dry Tube flow cytometry analysis to compare with the wet tube flow done by the NHS. The study was designed as presented in **figure 13** based on possibilities of recruiting samples (clinics not running as per before pandemic) while **Figure 14** illustrated the conventional wet tube FC workflow compared to DT.

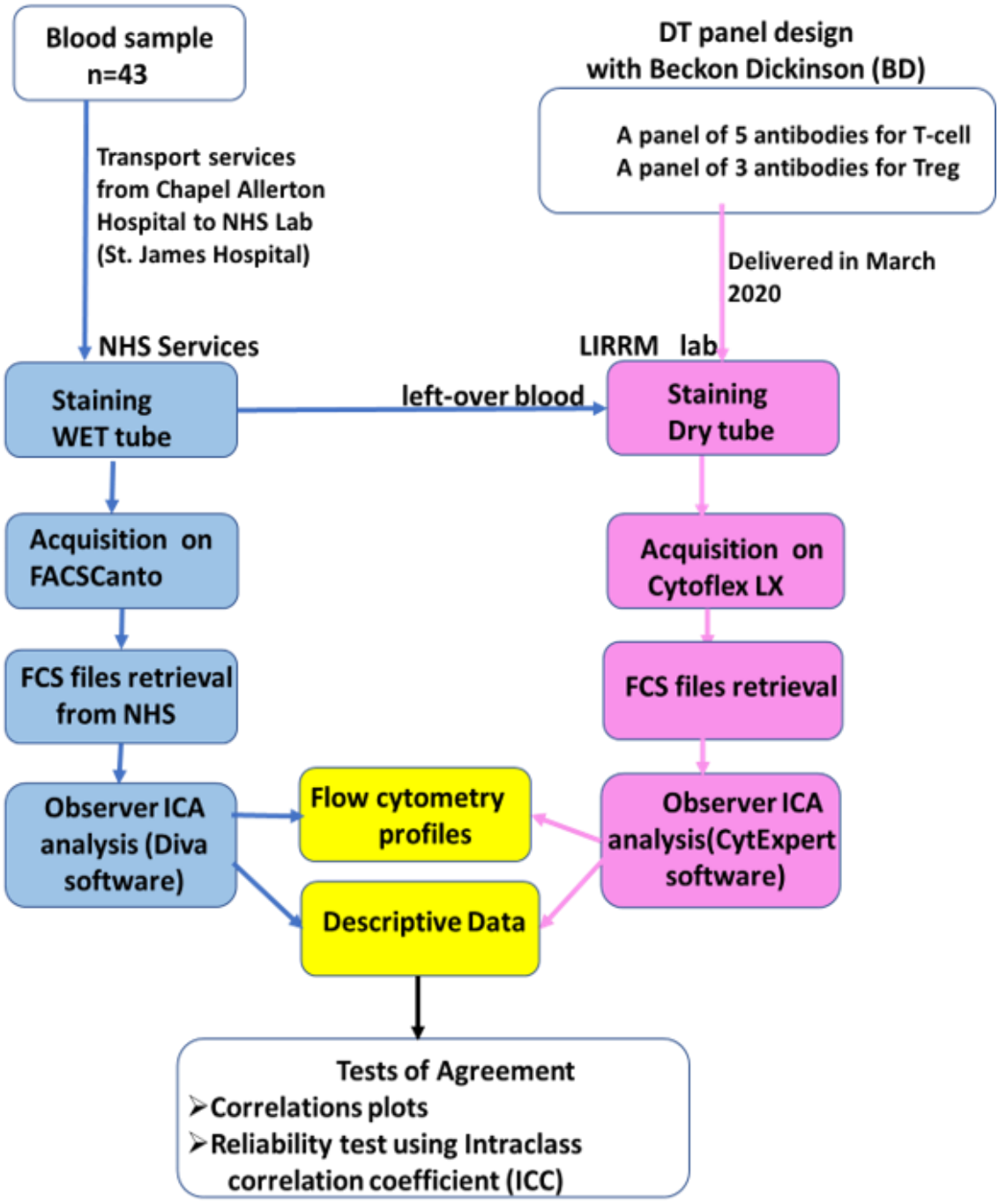


Figure 13 Flow chart of the study design

To assess the appropriateness of the DT technology, the study evaluated the performance of a T-cell panel (for identifying naïve/IRC subsets) and Treg panel against WT assay. Staining with WT and data acquisition on a Quanto machine and DIVA software was performed by NHS services for the samples (n=43). Blood left over from NHS lab were given to me on the day. Staining was performed with the DT and data were acquired on the Cytoflex LX/ CytExpert software. FCS files were retrieved and analysed later by ICA. Raw data (FCS files) of WT assay also were retrieved from NHS servers and analysed by ICA independently.

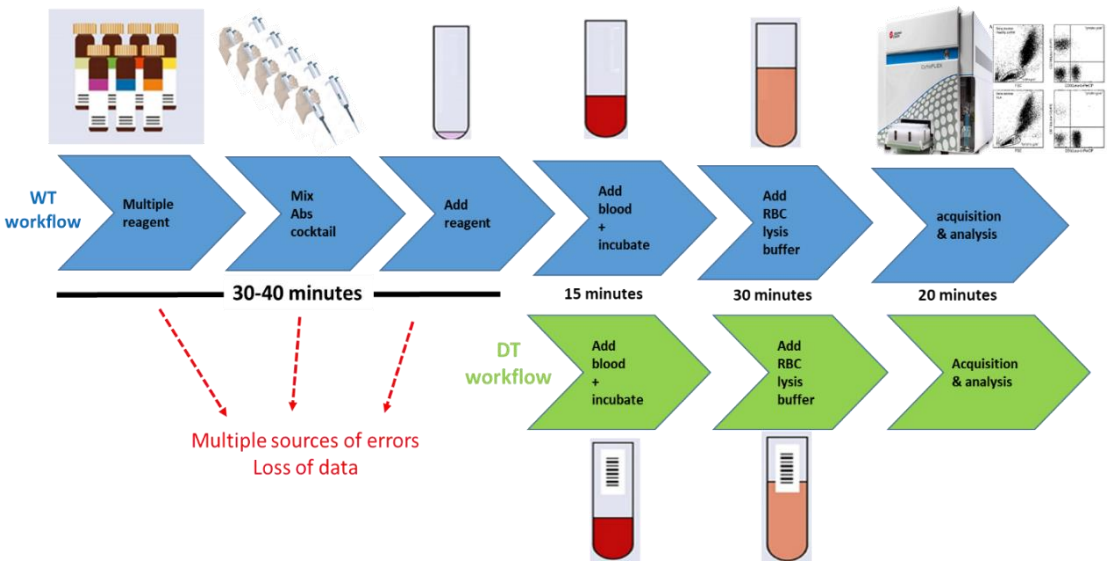


Figure 14 Flow cytometry workflow: Conventional vs simplified DT illustrated.

The workflow of conventional WT compared with a more simplified DT involves 6 steps (blue) with the first half of the procedure constituting the areas of multiple sources of error in flow cytometry analysis that can compromise the quality of data thereby leading to data loss whereas DT presented only a 3-step workflow (green) WT= Wet tube, DT=Dry Tube.

3.6.3.3 Dry tube cell surface and intracellular staining

In the absence of SOP for dry tube (DT) technology being a novel test development (see **figure 15** for a sample of DT kit and **table 18** for antibody cocktail used in the DT, (**appendices L and M)**, I adapted sample preparation from the NHS lab-processing SOPs for flow while accounting for the manufacturing instruction for dry tube.

Figure 15 Sample of novel DT KIT.
This shows the naïve T cell panel and Treg panel developed for the DT technology

Table 18 Dried antibody cocktails used in the DT Technology compared to antibodies used in the Wet Tube (WT)

DT			WT		
T-cell panel					
Specificity	**Format**	**clone**	Vol (ul)	**Format**	**clone**
CD3	PerCP-CY5.5	UCHT1	4	V500	UCHT1
CD4	PE-Cy7	RPA-T4	2	BV421	RPA-T4
CD62L	APC	DREG-56	2	APC	145/15
CD45RA	BV421	HI100	5	PE	Bio-rad MCA88PE
CD45RB	PE	MT4	7.5	FITC	Bio-rad MCA1114F
Treg panel					
CD3	Not used	-		V500	UCHT1
CD4	PE-Cy7	RPA-T4	2	BV421	RPA-T4
CD25	BV421	2A3	2	PE-CY7	2A3
CD127	Alexa Flour647	HIL-7R-M21	5	Per-Cy5.5	HIL-7R-M21
Foxp3*	Alexa488	236A/E7	5	Alexa488	236A/E7

*NOT contained in the Tube but added during the intracellular staining procedure.

Two hundred (200) µL blood was transferred to the lyophilized DT. The blood and the antibodies were mixed by inverting the tube. Incubation was carried out in the dark for 20 minutes at room Temperature (RT). After which eBioscience RBC Lysing solution was added directly to the tube (2mL/test sample) vortexed, and another of 2mL lysis buffer added. The mixture incubated at room temperature for 10 minutes in the dark.

After incubation, the tubes were spun at 500g for 5mins. Supernatant was discarded. Vortexing was done briefly to disperse pellet and 4ml of PBS wash buffer were added. Mixture was centrifuged at 500g for 5 minutes at 12°C and this washing step repeated twice. Cells were finally re-suspended in 300µl of FACS buffer and acquired on the CytoFlex Flow Cytometer.

For Treg tube, intracellular staining was performed using the BD Human FoxP3 Buffer Set kit. Re-suspended cells (at the end of the previous steps) were permeabilized using Fix/Perm buffer solution. Two millilitres (2mL) of buffer A (Fix) were added into the tubes and incubate for 10 minutes. The tube was centrifuged at 500g for 5 minutes, and the supernatant decanted. Cells were re-suspended by vortexing. After which 0.5 mL of buffer C (Permeabilization) were added, mixed by vortexing and incubated for 30 minutes at RT in the dark. After incubation, 3mL of PBS wash buffer were added and washing steps repeated as above twice. Cells were re-suspended by vortexing, 5µL of FOXP3-Alexa488 antibody were added, mixed and incubated for 30 minutes in the dark. The cells were washed as above, supernatant decanted and cells were finally re-suspended in 300µl of FACS buffer and acquired on the CytoFlex Flow Cytometer.

3.7 Flow cytometry analysis and gating strategies

Across the different study groups, gating strategies were similar as measuring the same subsets. Sample acquisition was performed on 13 Detectors, 4 Lasers CytoFLEX S VA-B4-R3-12 Flow cytometer (Beckman counter, USA). and analysed using Beckman Counter CytExpert software 2.4. Additional markers were added to the GOLMePsA study specifically. Ten thousand (10,000) events of interest were recorded for each subset (ie. >10,000 Treg rather than 10,000 CD4+T-cells) where possible. **Table 19** summarised the phenotype and subsets interrogated with their identification markers whereas the gating strategies used are presented in **figure 9.**

Lymphocyte subsets panel enumerates the different lymphocyte subsets (% of total lymphocytes) according to their expression of CD3, CD4, CD8, CD19, CD56 lineage markers.

- A measure of HLA-DR expression (%) was made on each lymphocyte subsets in the GOLMePsA study.

B cells panel: enumerate based on CD19+ cells, naive B cells (CD27-/CD38low), memory (CD27+/CD38low), and putative regulatory B-cells contained in the transitional T2 subset (T2/Breg CD27-/CD38high/CD24high).

- A measure of IL-6 receptor (IL-6R) expression was made on total B cells and subsets, by median fluorescence measurement (MFI) and HLA-DR expression (%) in the GOLMePsA study.

T cells panel: enumerated (as % of total CD4+ or CD8+ T-cells) naive (CD45RA+/CD62L+), memory (CD45RA-/CD62L-), and the IRC (CD45RA+/CD62L-).

- The expression of the IL-6R on individual subsets based on their MFI as well as their HLA-DR expression (%) in the GOLMeLPsA study.

Th17 cells panel: This was only performed in the GOLMeLPsA study only. Th17 cells (as % of CD4+T-cells) were enumerated using a validated surrogate phenotype CD161+/CCR6+/ CXCR3-.

- The expression of KIR3DL2 receptor measured (MFI) as well as HLA-DR % expression was recorded.

Table 19 Summary of markers defining subsets

	Lineage Markers	Subset markers	Additional markers	Subsets
Lymphocyte panel	CD3+CD4+ CD3+CD8+ CD19+ CD56+CD3- CD56+CD3+	 CD56^{high} CD56^{low}	HLA-DR	CD4+T-cells CD8+T-cells B-cells NK-cells (NK56^{bright}) (NK56^{dull}) NKT-cells and HLA-DR+ subset
B-cell panel	CD19+	CD27-, CD38- CD27+,CD38- CD24+, CD38+high CD27+, CD38+	IL-6R HLA-DR	Naive Memory Putative Breg Plasmablasts and IL6+ MFI/subsets HLA-DR+ subset
T-cell panel	CD3+CD4+ CD3+CD8+	CD45RA+CD62L+ CD45RA-CD62L- CD45RA+CD62L- CD45RA-CD62L- CD45RA-CD62L- CD45RA+CD62L-	IL-6R HLA-DR	NaiveCD4+ MemoryCD4+ IRCCD4+ NaiveCD8+ MemoryCD8+ IRCCD8+ and IL6+ MFI/subsets HLA-DR+/subset
Th17 panel	CD4+	CD161+CCR6+CXCR3-	KIR3DL2 HLA-DR	TH17 and KIR3DL2+ MFI HLA-DR+ subset

IRC: inflammation related cells. NK natural killer. IL-6R: interleukin-6 Receptor. Th17: T-helper 17. KIR: killer-cell immunoglobulin-like receptor. MFI mean fluorescence index. HLA-DR: Human leucocyte antigen- D receptor

3.8 Statistics analysis

All data were analysed using the SPSS V27.0 (IBM, Chicago, IL, USA). Figures were generated using GraphPad Prism 9.3.1 (La Jolla, CA, USA) and SPSS. R package was used for a specific function not available in SPSS with help from Dr Farag Shuweihdi (FS), one of my supervisors. P-values of less than 0.05 were considered significant.

3.8.1 Data description

Descriptions of numerical data are expressed as the median (interquartile range) and nominal data with numbers (%) where necessary (unless indicated otherwise).

3.8.2 Data normalisation test

Data normalisation tests were performed to investigate the normal distribution of data or otherwise using Q-Q plot. Non-parametric tests were used as most variables were not normally distributed.

3.8.3 Lymphocyte subset age-normalisation

Age-relationship for specific cell subsets were performed using a spearman correlation as previously described [4, 5]. Subsets with significant age-relationship ($rho>0.600$, $p<0.05$) were corrected based on patients' age using the correlation equation generated. Values obtained after age correction for patient data were termed normalised subsets frequencies as in previous publication [5]).

[Normalised subsets frequencies] = observed frequency [actual value obtained from analysis) – [expected frequency]. [Expected frequencies] = Correlation coefficient (β) x [age] + Constant(C) per subsets.

3.8.4 Univariate analysis of blood cell subset frequencies and for high/low risk of progression outcome

Exploration of data was performed in univariate analysis using Mann-Whitney U and Chi-square (X^2) tests to compare continuous and nominal variables respectively, between outcomes. Variables at baseline were then assessed individually for association with an outcome using unadjusted MWU to select subsets associated with the outcome and the area under the ROC curve (AUROC, 95%Cl) was also calculated as a parameter predictive of the outcome. Variables were also assessed individually for predictive value using unadjusted odd ratio OR (95%Cl) area under the ROC curve (AUC, 95%Cl). Sensitivity, Specificity, positive and negative predictive values (PPV/PNV) were calculated from the classification matrix of accurately predicted versus non-accurately predicted cases obtained from the probability of progression calculated by logistic regression. Bonferroni correction method for multiple testing [698, 699] was applied where relevant.

Statistically relevant cut-off value was obtained from the coordinates of the ROC curve at 80% specificity allowing dichotomisation for high/low risk. These were then used to establish sensitivity, specificity, PPV, and NPV of each subset for prediction of progression across the IAC. A spearman correlation-based clustering algorithm (Cluster-3, Stanford University 1998-99) was applied to assess collinearity between cell subsets and displayed using TreeView as heat-map.

3.8.5 Data missingness

All study cohorts had missing data points in LS from one panel or the other, but data were missing at random as confirmed with a Little's MCAR test ($p<0.0001$). Replacement of missing data were carried out using multiple imputations (5 cycles) method in SPSS 27. Given that the assumption of data 'missing at random' holds (Little's MCAR test, $p<0.0001$) and the study involves multiple variables, the method used provides better estimates that both preserve the underlying relationships among variables and of the reliability of estimates. In multiple imputation, known values of all variables are used to provide several sets of estimates of the missing data. Therefore, using this regression-based method the result produced estimates (i.e., regression coefficients and standard errors) that were unbiased with no loss of power. A pooled dataset resulting from the imputation process was utilised in subsequent analysing and modelling [700-702]

3.8.6 Multivariate regression model to predict progression across IAC

Logistic and Cox regression modelling was used to construct multi-parameter models for the prediction of outcomes across the IAC using a stepwise forward method, selecting the best combination of variables with the most significant improvement of the fit.

To assess the reliability of models, a bootstrapping technique using 500 permutations was developed and a Dxy discrimination index corrected for optimism was calculated in collaboration with Dr FS. Somers's rank correlation between the predicted and the actual outcome for both logistic and cox regression models were done using *R.4.1.3* (as not available in SPSS).

Individual probability to progress for each Cox model was used to check proportionality assumption. AUC for Cox regression were construct in R 4.1.3. Using statistically relevant cut-off value obtained from the coordinates of the ROC curve at 80% sensitivity, models performance accuracy in terms sensitivity, specificity, PPV, and NPV for prediction of progression across the IAC were also calculated from the classification matrix. WALD tests from the regression analyses performed were used to assess the contribution of each of our predictor variables.

3.8.7 Comparison analysis for validation of DT technology and NHS versus ICA (me) inter-observer variability analysis

Results are presented using correlation plots using data from the DT vs Wet tube (WT) to describe the strength of the relationship between both Methods. Bland-Altman plot using the difference between data from the DT and WT vs their mean for each CD4+T cell subset was constructed to access the degree of agreement by estimating the bias (the measurement difference) between DT and WT. The evidence of no bias between the methods/observers was performed using a one-sample t-test of the difference between methods or observers. Estimated reliability between the two methods/observers was also calculated for each CD4+T cell subset using intraclass correlations (ICC) (and its 95% confidence intervals (CI). ICC is a reliability parameter evaluating the correlation between duplicate analyses made either between 2 observers or by the same observer on 2 occasions on the same raw data. It takes values between 0 and 1, with a value of 1 corresponding to zero measurement error and a value of 0 indicating variability in analysis suggesting the presence of measurement error [703]. The reliability of the result between methods/observers was estimated with Intraclass Correlation Coefficient (ICC) (Type A) by the two-way mixed effects model where

people effects are random and measures effects are fixed and defined to measure for an absolute agreement based on the 95% confident interval.

3.8.8 Technical validation analysis of frozen samples in flow cytometry work

Correlation analysis between Fresh versus frozen samples was calculated using the non-parametric Spearman rank test to explore the nature of the relationship between two continuous variables. The comparison of slope values with a full identity (slope = 1) was performed after linear regression analysis using the F-test. Comparisons between fresh and Frozen PBMC were assessed in paired matched data using the paired t-test and Wilcoxon signed rank test (two-tailed). All statistical analyses were performed using GraphPad Prism version 10. Statistical significance was established at a p-value ≤ 0.05.

3.8.9 Trajectory clustering analysis

A series of cell subsets biomarkers across the LC panel, T-cell panel, B-cell panel, and Th17 panel were assessed at baseline, 6 months, and 12 months, representing variable trajectories. Data distribution across each panel was visualised using box plots.

Our objective was to identify hidden groups (clusters) by considering the collective trajectories, which would provide insights into the distinctive features of each cluster. To achieve this, we employed k-means clustering tailored for trajectory data. The analysis was conducted using the R package. Thanks to one of my supervisors, FS who supported me in doing this aspect of my analysis because data needed to be coded in R and I did not have time to learn the R package in year four of my PhD.

Exploration of data was performed for univariate analysis using Two-way ANOVA tests, between clusters A and B to determine the main effects of patients' cluster groups (1st independent variable) and time points: BL, 6M, and 12M (2nd independent variable), and their interaction effects on the frequency of cell subsets (dependent variables).

A Spearman correlation-based clustering algorithm (Cluster-3, Stanford University 1998-99) was applied to assess collinearity between cell subsets and displayed using Treeview as a heat map.

Results were displayed as mean (SD) for continuous variables. The P-value of less than 0.05 was considered significant.

Chapter 4 Result

4.1 Technical validation of the use of FC as a technology in biomarker research

Introduction

Flow cytometry is used a lot in research but has also made its way into clinical routine lab assessments. The technology is safe and robust but heavily depends on adherence to protocol and QA/QC while human data interpretation is also an integral part of its delivery.

4.1.1 Validation of the use of fresh blood flow cytometry as a biomarker technology

Blood samples from various studies were sent to the NHS immunology services, and LS data were acquired. An analysis of raw data from panels was carried out for a wider appreciation of the FC technology itself as a tool in clinical practice on one hand still being developed as a research biomarker on the other hand. The raw data from 5 panels were made available as electronic files (FCS files). Data analysis for the quantification of LS included panels for lineage count, B-cell subsets; CD4+T-cell subsets; CD8+T-cell subsets, and CD4+Treg.

The main purpose here, was to analyse the type of issues encountered, necessitating data to be excluded due to poor quality of samples or faulty processing (non-adherence to sample preparation protocols), faulty acquisition while panels not being performed at all was also accessed.

4.1.1.1 Technical validation from staining to data acquisition

Data from 2,911 raw FCS files were used to examine these issues.

Problems encountered

(1) Poor quality blood: This can be observed for any test performed routinely.

(a) An acceptable quality of sample for optimal flow cytometry is presented (**figure 16 A**) that allows clear identification of cell population of interest.

(b) Poor red cell lysis often results from biological characterisation of the blood such as very high viscosity, inflammatory related causes. It is impacting the ability of fluorochrome to work optimally (**Figure 16 A**). RA being an IA, this was to be expected in a FC test and normally not a major issue if the red cell lysis (RCL) is repeated as is often the case in research but not in NHS settings. As such, I observed many samples where RCL was far from optimal and the sample is almost impossible to read. This affects the FSC/SSC and can also affect marker detection.

(c) Marker expression sometimes was completely "blurred", and it was impossible to define gates but difficult to explain. It is noticeable that there were delays in sample transit from clinics (in Chapel Allerton Hospital) to the NHS lab (at St James hospital), via NHS transport services and a change of vehicle at the Leeds General Infirmary. This may have led to sample deterioration. Delay ≥24 hours are likely to impact cell viability [564], the forward/side scatter, and the expression of markers and may account for such poor samples.

2. Faulty processing: In most cases this involved non-adherence to sample preparation protocol or omission of an antibody **(Figure 16 B**). Of note, issues with non-availability of certain antibodies in a panel was used to explain such cases. This impacted mainly the Treg panel with ~8.5% loss of data due to missing antibodies (FoxP3 or CD127) but sometime seen also in other panels as well. Another

processing issue was poor red-cell lysis leading to erythrocyte contaminating the FSC/SSC region with cells and making it hard to gate the lymphocytes.

(3) Poor instrument set-up: Another issue was the lack of adjustment of instrument settings on a daily basis resulting in inability to "find" cells on forward/scatter on a sample of "poor quality" **(Figure 16A).** This caused uncertainty in the initial gating of lymphocytes although, marker expression usually allowed analyses of data. This is an issue observed mainly in routine analysis as protocol are set "in stone", settings being fixed and not allowing to adjust the FSC/SCC on a daily basis.

(4) Lack of compensation: Some data reported on the NHS result server, were not aligned to my analysis performed on the raw data and led to a lot of data discrepancies between NHS reports and my observations. However, this only impacted 1 of the panels. NHS data acquisition uses fully compensated protocols as standards. It appears that one protocol was damaged over the time and data were then acquired uncompensated and analysed/reported as such. This seems to have been occurring every time the software upgrade automatically while I could not understand why it affected only one of the panels. If not noticed by the operator, it allows data to be saved while no analysis can should be done without a fully re-compensating data. This significantly impacts the ability to analyse straight away (**Figure 16C).**

Another cause was change in protocols over the years (2013-2020) notably in the B-cell panel that initially used 8 antibodies but reduced to 5 two years later, while the protocol was not recompensated leading to over-compensation between certain fluorochromes. This was, however, a technical hurdle that could be overcome **(Figure 16D)** but resulted in a highly time-consuming process.

Comment reporting any of such issue should have been included in the NHS records made for such samples to allow critical decision making by the clinicians. However,

still being a research test, this was not yet included and may also lead to discrepancies between me and the NHS reports.

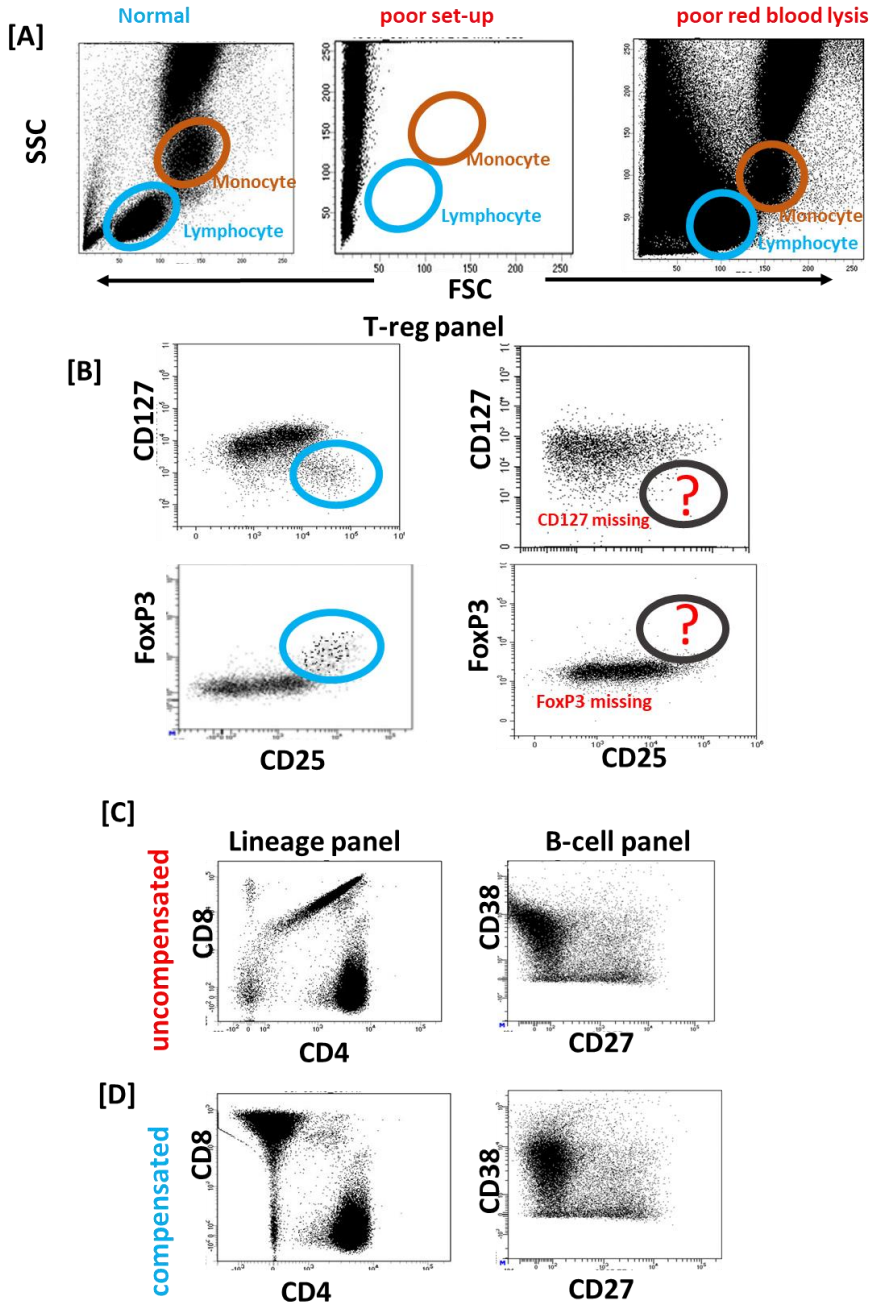


Figure 16 Example of FCS plots illustrating problem encountered during analysis.
(A) Example of displayed event in dual plot of size (FSC) versus granularity (SSC) of good (left) versus poor-quality of blood, poor instrument set-up (middle panel), or inefficient red cell lysis (right panel). (B) Plot showing missing of stains for CD127 and FoxP3 (black circle) likely due to the antibody not being added and plot (blue circle) represents Treg gate. Example of sample showing uncompensated plot in lineage and B-cell panel(C) followed by (D) the same re-compensated plots, post-acquisition.

Developing sample exclusion rule

Based on the observations, I developed 3 types of exclusion rules to allow results from acceptable QA/QC raw data to be selected for my study, considering "human" error, technical error, and poor sample quality.

I directly compared the 2 data sets and classified samples that needed exclusion based on the type of error eligibility associated with its exclusion

Processing error

Results showed that "human" error occurred more in the Treg panel analysis (8.45% of samples) compared to the addition of the same error in all other 3 panels together (Lineage, T-cells, and B-cells) at 1.61% of cases. **Figure 17** presented the outcomes of developing the 3 major exclusion rules for all the panels analysed. High level of multiple pipetting procedure associated with the Treg panel due to 2 levels of staining (surface and intracellular) may explain in part this observation although it is hard to explain why a panel whose gating depends on 1 particular marker (FoxP3) can be performed and reported when such antibody is not added (this occurred over 4 weeks of consecutive samples that were acquired and reported while the FoxP3 Ab was missing)

The second source of error was for Treg frequencies reported in the absence of CD127;

1) It leads to over-estimated Treg frequency and

2) This change should have been noted in the reported data provided to clinicians to make decision. The reason provided was the lack of commercial availability of this specific Ab clone. This is a serious issue that was reported and discussed with the NHS lab manager (who was unaware of it). Although it may have a limited

effect on data according to other studies where CD127 is not used [480] as it is a serious departure from the SOP in this context which included the CD3, CD4, CD25, CD127, and FoxP3 markers as the minimally required markers to define human Treg cells [704].

Other panels showed few faulty samples associated with processing error most likely genuine mistakes with pipetting.

Technical error

Technical errors involving faulty acquisition were either due to lack of compensation or arising from lack of adjustment in the instrument settings as described in the previous section. It was more pronounced in B-cell panel in 6.66% of samples

Poor blood quality

Poor blood quality appeared to affect the T-cell subsets panel more (2.58% of samples) impacting directly on the expression of the markers and rendering gating impossible.

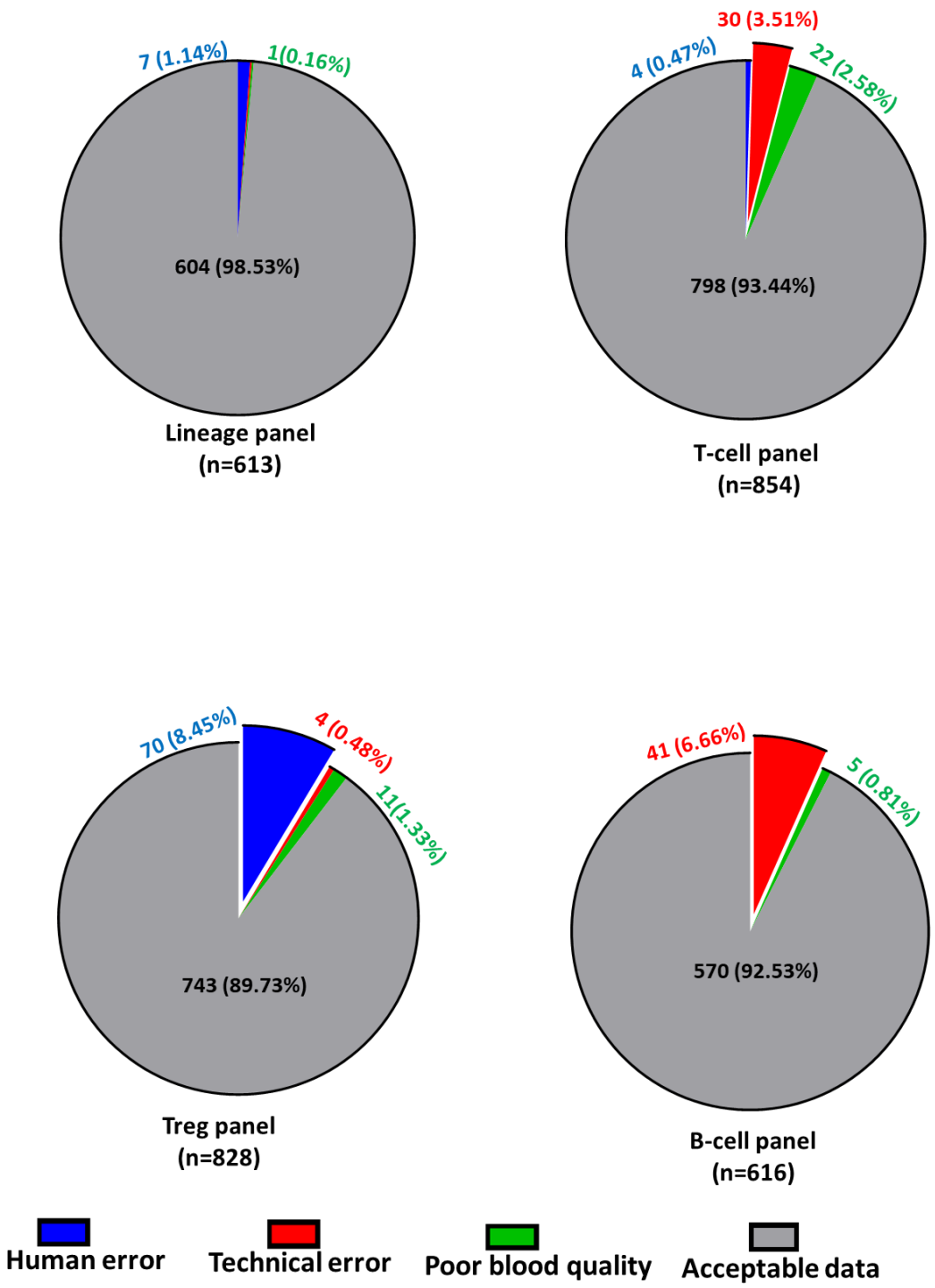


Figure 17 Exclusion rules for NHS flow cytometry raw data files available for the study.

Each pie chart represented the results of excluding samples based exclusion rules for sample fully reported on by the NHS flow cytometry services "processing" error (blue), technical error (red), and poor blood quality (green). Acceptable samples (passed QA/QC) are represented in grey colour.

4.1.1.2 Performance evaluation of inter-observers' variability in gating between routine NHS and myself (ICA)

Flow cytometry raw data from the 2 main panels (T-cell and Treg) routinely reported by NHS were re-analysed independently by me (ICA). The full SOPs used by the NHS including the gating processes designed by my supervisor originally when the protocol was transferred from her lab are attached as **appendix N.** The NHS-reported results for these patients were made available and utilized to determine the inter-observers' variability in the cell subset frequencies reported.

Paired samples (n=262 for naïve CD4+T cells, n=203 for IRC CD4+T cells, and n=162 for Treg CD4+T cells) with results available from both observers (NHS versus ICA) were included in this analysis while I excluded data for samples fulfilling my rules for exclusion defined above.

Distribution characteristics between observers

I was first trained by a flow cytometry expert with over 20 years of experience (Dr Ponchel) who also developed those panels in research setting as detailed in the method section. I then compared my analytical skill to her and when deemed “trained”, I performed my analysis of these raw data files.

To validate the reproducibility of data between observers I considered myself as the reference and the NHS-reported data as the challenger.

NHS results being acquired by various numbers of staff over 2013-2022, were divided into 3-time period representing (as close as possible) the 3 period of employment of these staff. The general characteristics of the distribution of data for the three blood cell subsets (Naive, IRC, and Treg CD4+T cells) between

observers (Me and NHS2013-15, NHS2016-19, and NHS2020-22) were evaluated using paired t-test.

Results are presented in violin plots **(Figure 18)** comparing 3 periods of time corresponding 1st to the development of the service from the right panels, with staff 1 (2013-2015), 2nd , to routine use of the service with staff 2 (2016-2019), and 3rd over COVID pandemic with staff 3 (2020-2022) and my analyses on the right.

For Naïve, the mean value between the two observers at 2013-2015 did not differ significantly (40.30 vs 38.64, p=0.6809) with a very negligible mean of difference (95%CL) [0.1667(-0.6540 to 0.9874)]. From 2016-2019 period, the data distribution showed a significant decrease in NHS data (36.94) compared to mine (39.72, p<0.0001) with a mean of difference 4.392 (1.309 to 2.897). This suggested a loss of data consistency over this period. In 2020-2022, there was no significant difference between the two observers (33.14 vs 34.27, p=0.0600). The mean of difference of 1.249 (0.0265 to 2.471) seems to indicate a restoration towards consistency. This may likely be due to a shortage of staff available during the pandemic compared to 2016-2019 when several experienced staff had been involved. The overall correlation coefficient (r > 0.9) for each period showed that pairing was significantly effective, suggesting that naïve data distribution is closely comparable between NHS and me.

For IRC, data distribution, the NHS across the 3 periods of time respectively did not differ significantly from data obtained by ICA (all p>0.05), suggesting a better consistency in overall gating pattern for IRC subsets.

For Treg, NHS2013-15 results were not significantly different between the two observers (p=0.8212) during the development of the service while during the late period (2019-2022) data were less closely comparable. This may suggest inconsistency in gating over the last two periods. However, this also included a lot of sample reports when the CD127 or FoxP3 Ab was missing. This highlights the need to develop reporting notes for issues that may affect data as well as maybe defining rules for not performing an assay when reagents are not all available (when NO FoxP3 is available with a caution note while maybe for when CD127 is unavailable).

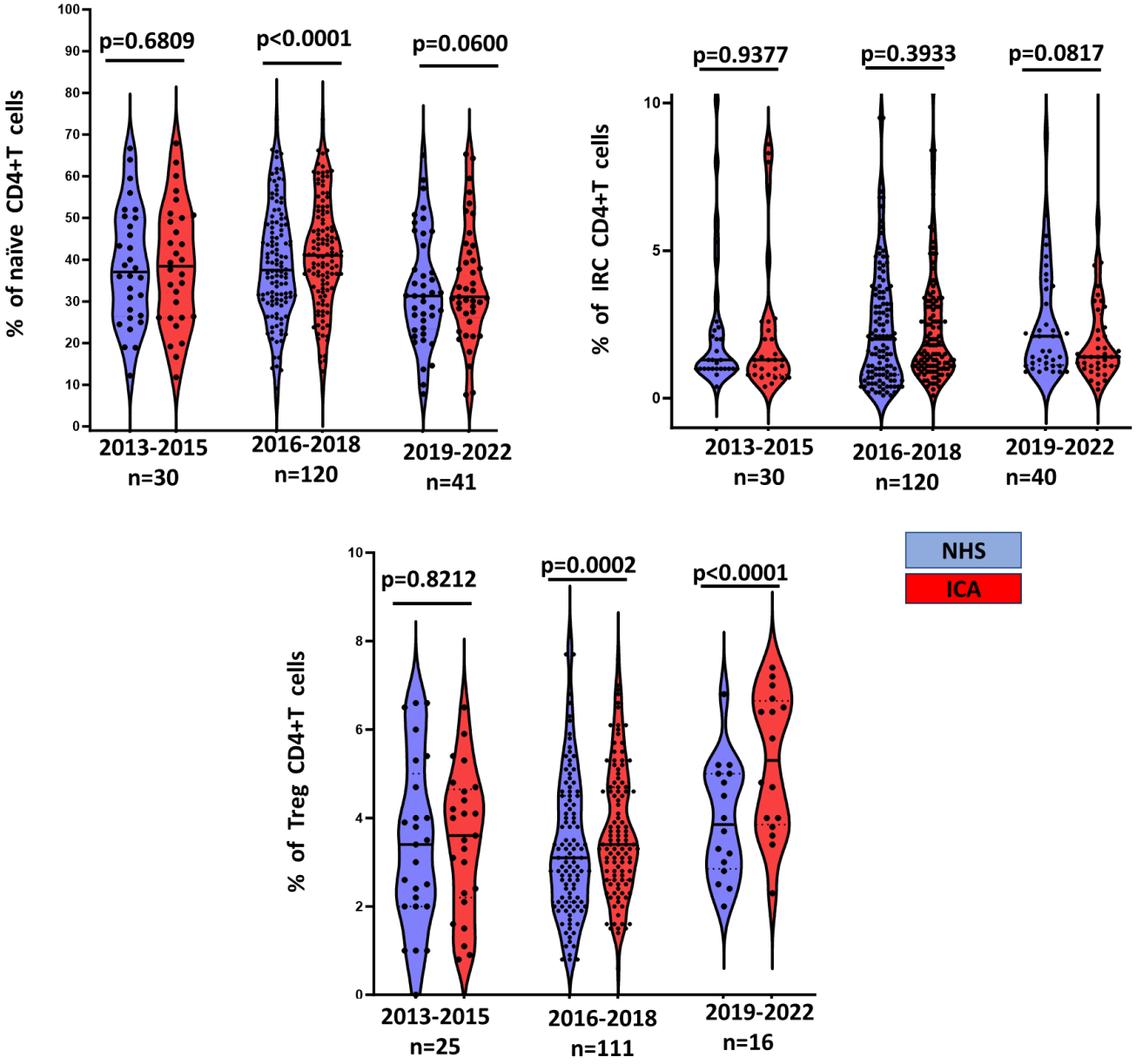


Figure 18 Distribution of results for 3 subsets gated by different observers. Violin plot indicating data distribution between NHS (blue) and me (red), thick line within the violin plot represents the mean value. A paired T-test p-value is indicated.

Overall correlation coefficient ($r > 0.6$) across each period compared showed that pairing was significantly effective, suggesting that data at the cohort level are relatively similar. I, therefore, decided to use all 3 staff/period as if a single observer and proceeded to a more detailed analysis.

Correlation between NHS data versus ICA

I then performed an analysis of the agreement between observers first using spearman correlation analysis. Paired data (n=262 for naïve CD4+T cell, n=203 for IRC CD4+T cells and n=162 for Treg CD4+T cells) between both observers (NHS versus ICA) were included in the analysis.

Results are displayed using a simple scatter plot with the line of best-fit to address whether it is a perfect match compared the line of equality which would suggest an absolute correlation between datapoints between the two observers **(Figure 19).** The correlation coefficients (rho) indicate the strength of the linear relationship between two variables

For naïve CD4+T-cells, there was a highly significant correlation (rho=0.9336, $p < 0.0001$) and both the line of best fit and the line of equality were closely overlapped.

For IRC CD4+T, the correlation was still significant but with a decline in correlation strength (rho= 0.7568, $p< 0.001$). The line of best fit showed divergence from the line of equality with increased IRC frequencies in samples particularly at frequencies above 5% in my analysis. This shows a poor linear relationship between observers when frequencies are low (<5%) while the higher frequencies are actually driving the correlation. This suggests a poor consensus between observers.

For Treg, the correlation coefficient reduces even further (rho= 0.6877) showing a decline in the strength of the relationship between the datasets. Lines of best-fit and equality were still relatively close. Several points were particularly different. Two points notably suggested a possible additional sources of error in data entry on NHS server as reporting very high values (9% and 10% frequencies). Further analysis was re-done excluding those odd points. This suggests a particular need for experienced staff to gate this panel and possibly as well the need to account for data point where CD127 is missing.

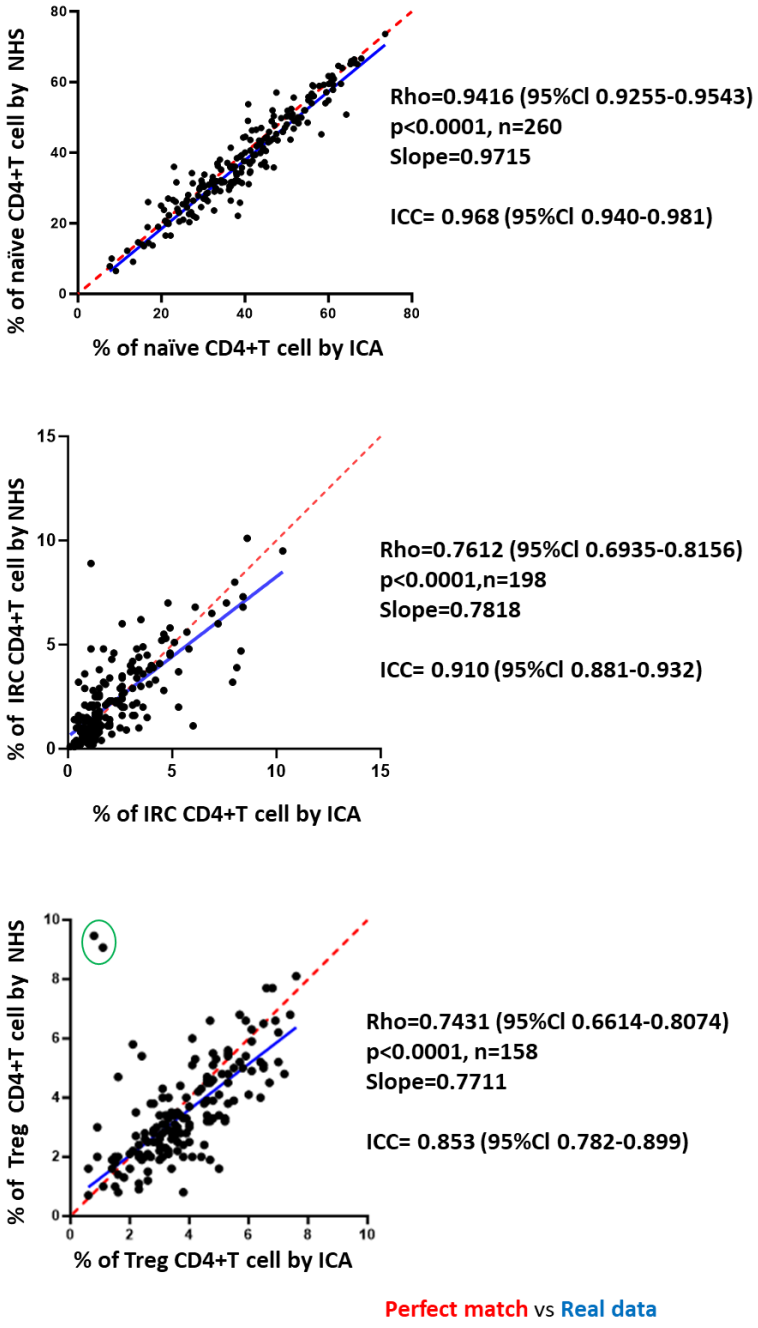


Figure 19 Comparison between frequencies identified by the NHS and the ICA (Spearman correlation analysis).

Data displayed in scatter plot with Line of best fit (dotted line, slope), line of equality (blue line, slope=1) and rho = coefficient of correlation indicated in for each subset. Intraclass correlation coefficient (ICC) indicates the reliability results between observers

As a second step to assess level of the agreement between observers and explore further the behaviour of data distribution, Bland-Altman (BA) plots were constructed for each cell subsets by plotting the difference between measurements from the two observers (NHS - ICA) against the average of both observers **(Figure 20)**. If no difference were to be seen, the mean of those differences (green) should be equal to zero (black line) and the 95% CI should be small (dotted red lines).

For naïve cells, the result of the mean differences (measurement bias between the 2 observers) was approximately –2 which is close to zero as naïve cells frequencies are usually about 30%. The BA plot suggests a range in the difference between the 2 observers from -10.5 to +6.3 (with 95% CI) **(Figure 20** top panel). Thus, this 95%CI suggest a trend for higher naïve cells in IAC compared to NHS.

I then tested the hypothesis that the observed mean of the difference (1.9%) was indeed not different from zero (for evidence of no bias between observers) by performing a one-sample t-test of the difference . A statistically significant p value 0.001 was obtained which confirms a certain level of measurements bias (over estimation) between naïve frequency quantified by NHS and those quantified by me.

For IRC, the mean difference (green line) was superimposed on assumed mean difference of zero (black line) with little or no suggestion of bias. The result of the mean differences was approximately –0.01% which showed a non-statistically significant different when evidence of bias was testing using one sample t-test (p=0.400).

However, a difference in data distribution patterns was observed with data points tending to diverge farther away from the actual mean of zero at frequencies above 5%, suggesting possibility of increase in gating variability with increasing IRC frequencies. With the BA plot showing a range in measurement difference (-2.35 to +2.36) between observers that is often higher than the frequency of IRC themselves, it suggests that IRC results obtained by NHS are very poorly comparable to those obtained by ICA particularly when below 2% frequency. This will need to be considered when interpreting data for their biological value.

For Treg, the bias was –0.39 (green line) with a range within 1.76 and –2.49. However, the test of evidence of bias between observers suggested that the difference was significant ($p<0.001$). Interestingly, the data distribution pattern showed that the majority of the datapoints are scattered below the actual mean of zero (black line) suggesting a negative bias. This mean that, most result by NHS are under-evaluated compared to me.

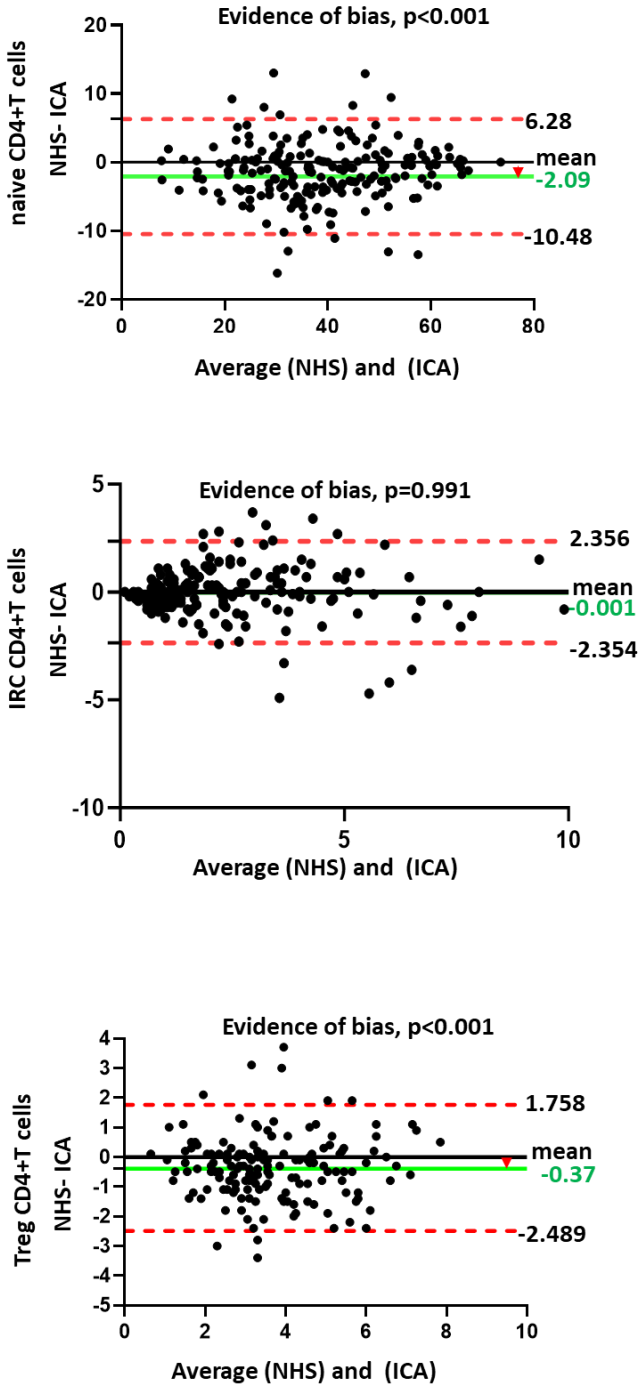


Figure 20 Bland-Altman plots of the CD4+T cells subset quantification comparing the NHS observer and ICA.
Data displayed in BA plot with range of results between observers (dotted red line), expected mean difference (dark line at zero), observed mean difference (green line). p-value for evidenced of biased measured by one sample t-test displayed

This BA analysis confirmed that spearman correlation analysis was not sufficient to fully acknowledge the probability of error between measurements between 2 observers as it does not measure actual bias or estimate their level of agreement (as represented in BA plots). As such, intraclass correlation coefficient (ICC), is recommended as a better index for reliability test between 2 measurements , and therefore was also performed as it reflects not only the degree of correlation (i.e., the strength of linear relationship) but also the agreements between 2 measurements [703] . High reliability or low measurements error should be equal (or close) to a score of 1 whereas low reliability or high measurements error will show a score close to 0. ICC values above 0.9 are considered of excellent reliability between 2 measures [705].

The average ICC (displayed on **figure 19, page 165**) between NHS and ICA was quite high for naïve cells [(0.963 (95%CI, 0.937-0.977)], indicating good reliability between both observers. For IRC [0.776, 95%CI 0.705-0.830] and Treg [0.781, 95%CI 0.678-0.848] scores were lower confirming discrepancy in data reported by NHS from that generated by myself. The involvement of different staff over the years could have played an important role in impacting gating consistency between NHS and me considering the subjectivity of the process and the need for experience in doing it.

Taken together, this data prompted me to re-do all the analysis of all raw data files used in my thesis rather than accept reported values by the NHS. This resulted in a highly time-consuming approach to my PhD data acquisition notably with the burden due to the need to recompensate all B-cell panels.

4.1.2 Flow cytometry validation using Dry tubes

The Dry tube (DT) project was driven by the need to improve quality and reproducibility of flow cytometry-based biomarkers while reducing the costs/time of the wet tube (WT) workflow in health services flow cytometry laboratories. This is an industrial collaboration with Becton Dickinson (BD) on the backbone of the biomarker research program funded by LIRMM. The DT is intended to minimise lab procedures (i.e., handling mistake, missing Ab, etc…) and gain time when used in routine settings and for clinical trials (i.e., like TEEMS) as well as to disseminate the test internationally.

The DT for the T-cell panel and Treg panel were designed for quantifying naïve-CD4+T, inflammatory related cells-CD4+T (IRC-CD4+T cell) and CD4+Treg which are the 3 most important blood cell subsets biomarkers used in RA patients. Although, BD was not certain that the intra cellular FOxP3 staining would be possible.

The DT antibody selection and optimisation were performed between my lead supervisor, Dr Ponchel and BD in 2019. Antibodies were dried in a 12 x 75mm tube for T-cell panel for CD4+T subsets and a 3- antibody cocktail for Treg panel. The antibodies for both panels are shown in **table 18, page 141)**. A total of 50 tubes were manufactured by BD with an expiration date 31/10/2020, delivered to Leeds in March 2020, a week before lock-down. Progress was therefore interrupted by the pandemic, precisely when the project had reached the step of performing a WT/DT data comparison on actual patient samples.

Data to be acquired were subsets % with no need for MFI hence the fluorochromes were adapted by BD based on technical & commercial considerations while the Ab clones were chosen after getting closely comparable staining pattern/spearman correlation results (rho>0.9) during the optimisation phases (see **appendix A**). The DT were manufactured using the same concentrations of antibodies that were used with the WT reagent, which was established in the prior validation/optimisation phase of the project.

The numerical quantification of these 3 CD4+T subsets (naïve, IRC and Treg) in samples from rheumatology clinics directly compared results obtained by the NHS immunology flow cytometry service using WT and by the LIRMM lab using the alternative clone and fluorochrome (performed by Dr Ponchel before my arrival).

The immunophenotyping investigation using the DT was evaluated in 42 patients for various LIRMM clinics, consecutively included in the study over a period of 3 months post pandemic. These patients came from different clinical studies at varied stage of the inflammatory arthritis continuum (IAC) and were analysed in parallel but independently to compare data obtained from the WT and DT assays for the enumeration of the 3 cell subsets.

Due to covid restricted access to NHS, this part of my thesis was delayed, and I had to acquire results in the LIRMM university lab (between June 2022 and August 2022) rather than in NHS service. Samples were received and processed by the NHS as usual. The residual blood was given to me to proceed for the DT. Flow data acquired by NHS services (FCS files) for WT data were retrieved and were compared with similar FCS files obtained from the DT assays in the university laboratory. This implicated the use of two different flow cytometers

complicating the DT data comparison as depicted in FC workflow **(Figure 13, page 138).**

4.1.2.1 Staining results

Because of the 2 years delay since manufacturing the DT and the expiration date on the DT being passed, I first assessed whether the phenotype of cells stained with the DT was satisfactory for the enumeration of cell populations not compromising the separation of any T-cell subset populations compared to sample stained with WT assay. As shown in **figure 21 (top plot,)** the use of a dried antibody cocktail in different fluorophores has limited impact on the detection of the naïve and IRC cell subsets, although the profile was not absolutely identical, particularly for the IRC. CD45RA staining showed better separation with the BV421 fluorophore conjugate in DT than in the WT with PE conjugate (left vs right column). The DT panel design seems to have enabled a better separation of IRC than the WT although the use of different flow cytometer machine for the two methods may have played a role as well (left vs right column).

For the Treg panel, the staining was not detectable for FoxP3 in DT (even with double antibody dose) compared to WT **(Figure 21, bottom plot**).The same antibody used in WT assay was however working fine suggesting that the Ab itself is not the issue but more likely the intra-cellular procedure not being compatible with the DT technology.

Despite different fluorochrome/clone conjugates in DT vs WT for surface proteins (CD4/CD127/CD25), a “putative” Treg phenotype (using only cell surface markers $CD25^{high}CD127^{low}$) showed similar staining pattern as well as similar frequencies (data not shown, n=17), confirming that surface staining is of good

comparable quality in DT vs WT. Further Treg data acquisition with DT was not undertaken as staining for FoxP3- was inconclusive.

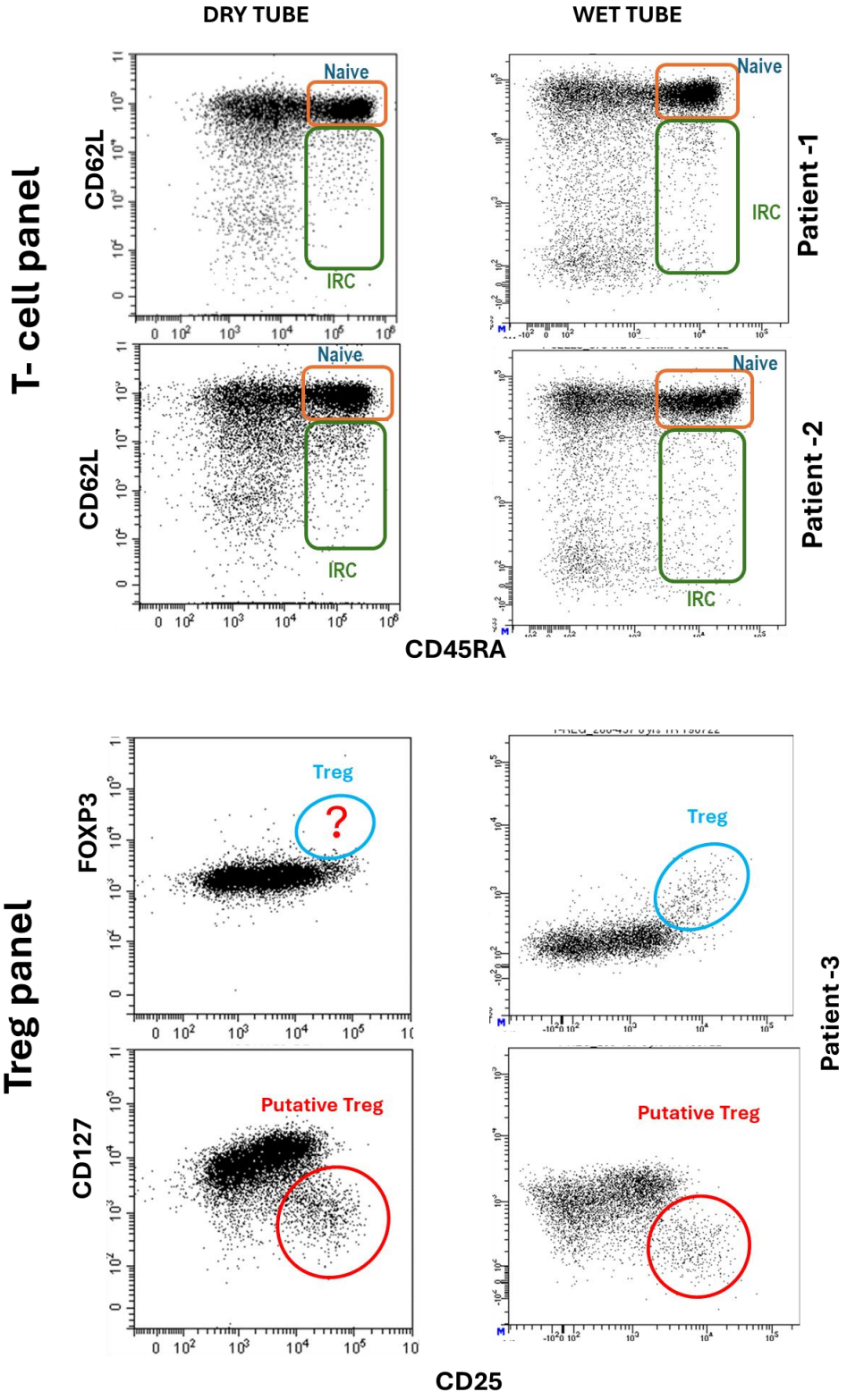


Figure 21 Representative flow cytometry profiles of whole blood samples stained with the DT (left column) vs WT (right column) technologies.

The dual plots shown are derived from the acquisition at least 10,000 events in a CD3+/CD4+ T cell gate, followed by gating for specific phenotypes. Plots represent the same samples acquired and analysed in parallel by the two Methods. For data acquisition/analysis, a Cytoflex LX /CytExpert software for DT and a FACSCanto /DIVA software for WT were used and plots from both methods were shown in biexponential displayed. For the T-cell panel, the figure shows 2 representative samples in patient 1 and patient 2. The gating strategy for the identification of $CD3^{+}CD4+$ T-cells with a naïve phenotype used $CD45RA^{+}CD62L^{+}$ (yellow square) and for IRC as $CD45RA^{+}CD62L^{-}$ (green square) suggesting good reproducibility. For the Treg panel (patient 3), CD4+T-cells were gated first. The FoxP3 staining was not observed in DT. A 2-colour gating strategy for the identification of "putative" Treg ($CD127^{low}CD25^{low}$ red circle) was used (bottom plot) and showed reproducible data. Treg could be identified in WT as ($CD127^{low}CD25^{high}FoxP3^{+}$ blue circle), while no convincing staining in the DT would allow for a clear population to be detected.

4.1.2.2 Performance evaluation of DT versus WT technologies

Data obtained from each method were then analysed using the same logic as for observers' comparison (section **4.1.1.2, page 159**), here comparing methods.

Data first were described using the median (interquartile range, IQR) for % frequency of the subsets numerated **(Figure 22A).** The naive median 33.15 (IQR 22.92 to 41.45) for DT did not differ significantly from that with WT (31.30 (25.10 to 43.40), p=0.8833

The naïve and IRC results were then analysed (n=41).

The agreement between the DT vs WT (scatter plot) showed the line of best fit being close to the line of equality **(Figure 22B)** for naïve. The comparison was significantly satisfactory (rho=93%, p<0.0001).

Bland-Altman (BA) plots were also generated for each cell subsets and Intraclass correlation coefficient (ICC) were calculated **(Figure 22C).**

For naïve cells, BA plot showed little or no suggestion of a bias between the WT and DT. The mean of the differences was relatively close to a zero (+0.8). The BA plot suggests a range in the difference between the 2 techniques from -7.5 to +9.1 (with 95% CI) indicating a range within which measurements differ for naïve CD4+T-cells enumeration between WT and DT similar to that observed between observers NHS vs me I then tested the hypothesis that the true mean of the difference was indeed zero. Non-significant p value of 0.225 was obtained, confirming no evidence of measurement bias.

ICC between naïve data obtained with both techniques was 0.971(0.947-0.985), also indicating good reliability between both methods for naïve enumeration.

For IRC, enumeration, the DT median was significantly higher 4.78 (2.69 to 7.32) than with the WT, 1.75 (1.10 to 3.50), p=0.0101) suggesting a difference in results obtained between technologies (**Figure 22A),** possibly due to the better subsets separation observed with the DT for BV421 CD62L Ab.

For IRC enumeration, the median difference was different from assumed mean difference of zero, with a more pronounced bias at –1.29%. A difference in data distribution pattern was observed with data points tending to scatter above the median line at low frequencies but below the median line at high frequencies, suggesting constant and/or proportional systematic errors present at high frequencies. The observation that DT enumeration of IRC gives a higher % result than WT is likely to be due to the change in fluorochrome and changes in clones showing a better separation in DT than in WT **(Figure 22C).**

The BA plot suggests a high range in measurement difference (-6.9 to +4.3) between WT and DT, with a definite bias (p=0.006) mainly observed for high frequencies (>2.5%) as over-estimated by the DT while at low frequencies (<2.5%) they were less affected. Reliability remained high over the overall range, despite better determination with DT for high frequency with an ICC=0.9629 (0.917-0.981) for IRC. This suggests that IRC results obtained by WT may likely be reproducible with DT technology.

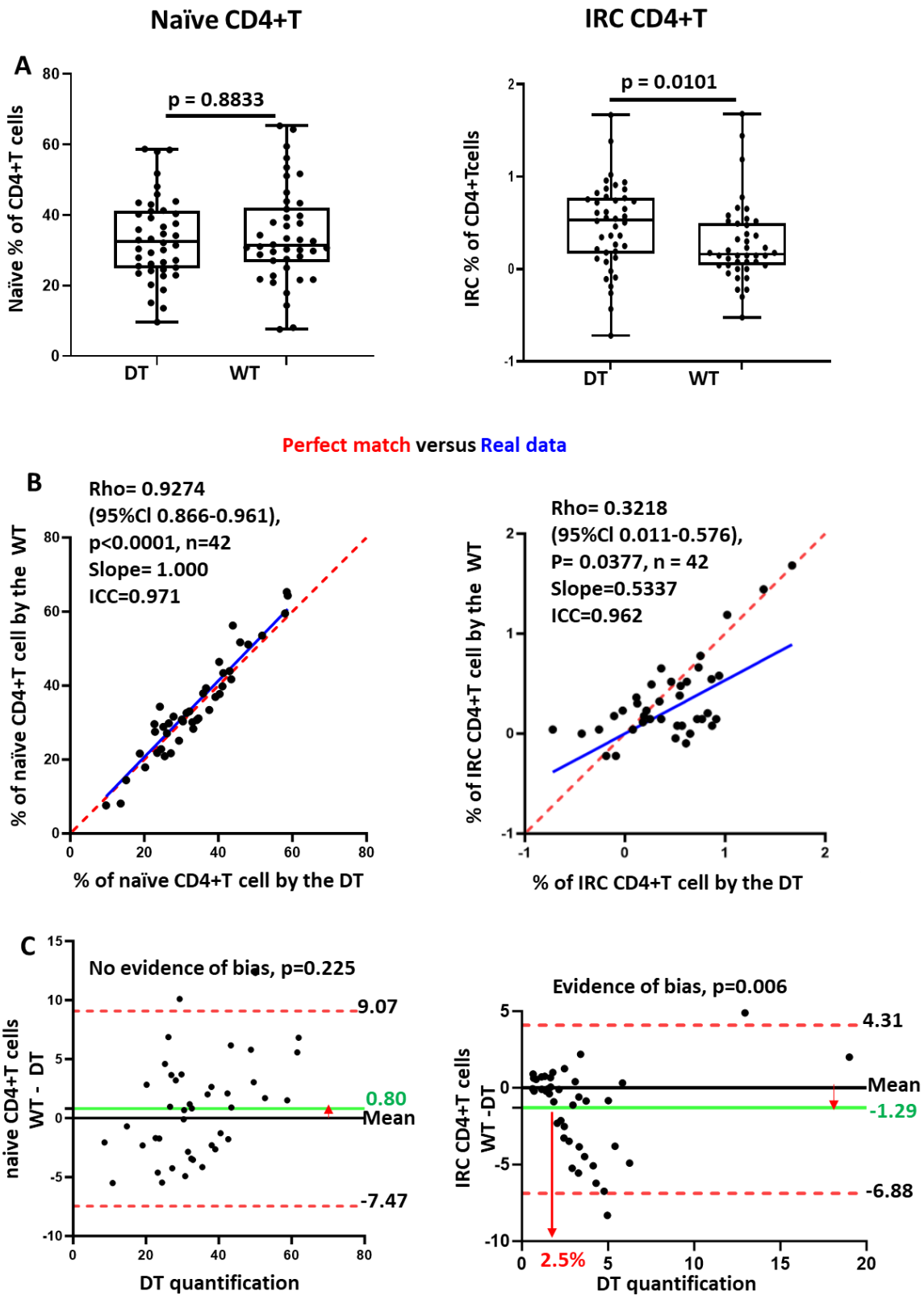


Figure 22 Data characteristics of subsets biomarker analysed (n=41)

(A) **Data distribution between DT vs WT.**Boxplot presented the data distribution between dry tube (DT box) vs wet tune (WT box). Mann Whitney U test, p values displayed.**(B)Comparison plots of the cell population frequencies identified with the WT and the DT**. Data displayed in scatter plot with Line of best fit (dotted line, slope), line of equality (blue line, slope=1) and rho = coefficient of correlation indicated for each subset. Intraclass correlation coefficient (ICC) indicates the reliability results between observers **(C): Bland-Altman plots of the subset frequencies identified comparing the WT and DT technology**. Data displayed in BA plot with range of results between observers (dotted red line), expected mean difference (dark line at zero), observed mean difference (green line). p-value for evidenced of biased measured by one sample t-test displayed

4.1.3 Flow cytometry validation on frozen samples

Immunophenotyping analyses of human peripheral blood mononuclear cells (PBMCs) using flow cytometry technique can be performed either on fresh blood samples as well as on frozen samples. Fresh samples are commonly preferred for routine analysis given the potential deleterious effect of freezing on functional and phenotypic properties of cells that could be associated with cryopreserved PMBCs samples [706-711]. However, in many research and clinical trials, cryopreservation has become a standard procedure for storage of PBMC and a very useful practice in multi-centre studies where blood samples must be collected for immune monitoring programs, therefore allowing for frozen PBMCs to be processed under identical FC conditions after transfer to centralised laboratories for FC analyses. This procedure allows for batch analysis of samples to enhance consistency which reduces the potential inter-assay[712] and inter-laboratory variation [713] which is often an issue in multi-centre clinical trials with local sample collection [714, 715]. As a drawback, it is highly possible that certain cell types and their surface markers may likely be affected by the cryopreservation of PBMC [715, 716].

Here frozen PBMCs samples (n=131) from the GOLMePsA clinical trial were analysed to search for potential associations between immune cell phenotypes and trials outcomes. PMBCs were isolated from freshly collected blood samples and processed at the LIRMM Chapel Allerton tissue bank (using LIRMM SOP 30S for the isolation and freezing of PBMC using Leucosep tubes, see **appendix E** for detail. Briefly, PBMCs were isolated from whole blood collected in Lithium Heparin tube (green top) . The lymphocytes were frozen in fetal calf serum (FCS) containing 10% dimethylsulfoxide (DMSO) for long-term storage (-150^{0}C). The frozen PBMCs were stored for a mean period of 3.28 years (95% CI 2.99 to 3.56)

due to the pandemic, after which they were transferred to LIRMM lab at St James hospital for FC analyses.

Panels were designed for cell surface markers/ cell phenotypes based on data available from a fresh samples study performed by TRIMID group [717]. Investigating any effect of freezing PMBCs on the FC results was essential for the quality of the analysis reported for this clinical trial. Therefore, I evaluated the effect of cryopreservation on the quality of FC data produced.

4.1.3.1 Frozen PMBCs were associated with reduced cell viability

To assess the value or rather possible limitations of using frozen blood samples in FC analyses, cell viability is one of the main issues affecting data quality.

The use of viability dye (such as propidium iodide (PI), 7-aminoactinomycin D (7-AAD), dyes which can selectively label dead cells with compromised membrane integrity) would have been ideal solution for discriminating between live cells and dead cells prior to flow analysis [718-720]. However, this was not included in the Trial protocol. I was restricted for consistency between the Trial protocol and the circumstance and time required to perform this analysis within 6 months while ethical permission to change the protocol would have taken too long and not a priority for the Trial team, thereby leading to the use of an alternative strategy here which provided assessment of live cell. Viability was defined as the percentage of cells from the total event of cell gated using a dual plot of FSC-width versus FSC-height which itself double gated from the plot of SSC versus FSC to separate viable from dead cells based on morphology **(Figure 23A**). Cell debris has reduced light scattering, and as such also need to be excluded giving an estimated % of live cells with this strategy.

I then used the same strategy **(Figure 23A)** to compare % of viable cells between fresh and frozen samples and the result presented in **figure 23B**.

The result of this analysis showed very large variations in the distribution of percentages of viable cells used for analysis across samples ranging from <1% to 69.3% in frozen PBMC (n=131, median, IQR), 45.30 (21.50-52.00) compared to fresh sample (n=37, 64.30 (58.30-72.50), MWU test p<0.0001

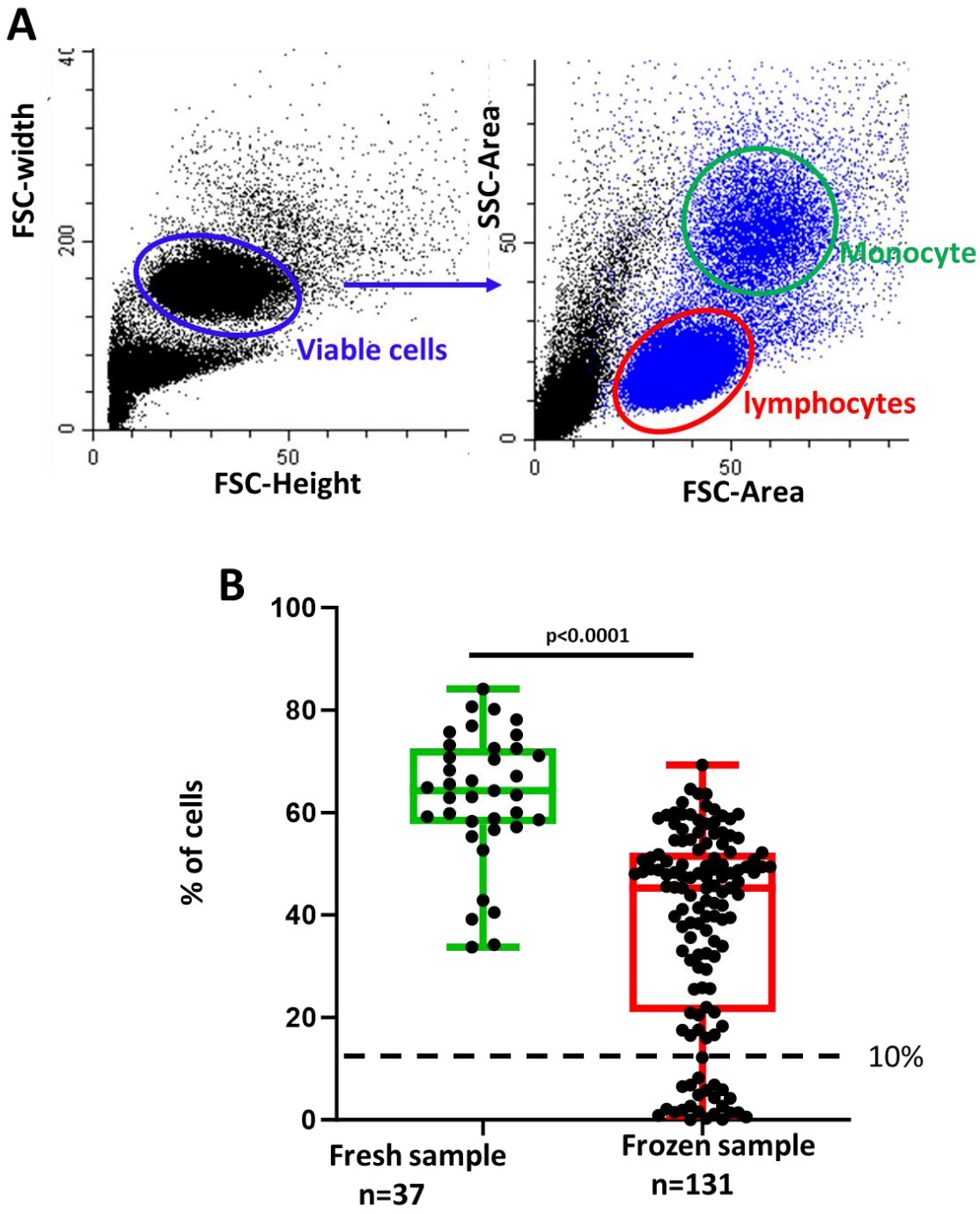


Figure 23 Gating strategy used to quantify viable cell fresh and frozen sample.

A. Gating strategy. The viable cell gate (blue) interrogated in the FSC/SSC plot (blue events) showing the lymphocyte gate(red) and monocyte gate (green).**B Comparison of % of cells selected for analysis in fresh sample(n=37) and for viable cells in frozen PBMC(n=131)**. Boxplot presented the data distribution showing a significant difference between Fresh (green box) vs frozen (red box). Mann Whitney U test, p<0.0001 displayed. The dotted line indicated cut-off % viability (10%) for possible exclusion.

I then classified the viability in frozen PBMC using percentage bins (decades of % windows) to evaluate which sample would be acceptable for further analysis. Based on this boxplot (**Figure 24A**), ¼ of the sample would be lost to the study at below 20% viability. Choosing a cut-off at 10%, 19% of the sample would need to be excluded.

To determine whether the loss of cell viability could be associated with the duration of storage, a correlation analysis was performed (**Figure 24B**, spearman test). Contrary to expectation, there was no significant relationship between cell viability (%) and storage duration in days (rho=0.03791, p= 0.6661). This suggests that other aspects of the freezing process may provide an explanation for the reduction in cell viability.

Further analysis was performed to determine if loss of viability is associated with inflammation levels (measured by C-reactive protein, CRP) which contribute to poor quality of cells FSC/SSC on fresh blood samples[721, 722] Sixty-three (63) samples were manually retrieved with the help of another PhD student for CRP level due to MHRA restriction in accessing clinical data from the trial electronic records. The result showed no significant relationship between CRP and cell viability in frozen PBMC either (rho= 0.1684, p=0.1870, **Figure 24C**).

Another factor could be the processing of samples (ie., the lymphoprep procedure). This has happened over 5 years while involving only 1 staff, with long experience of this procedure that is furthermore regulated by an SOP as was the defrosting process therefore ensuing all samples were processed the same. Furthermore, when a few samples were processed on the same day from several patients/time points, differences in viability were still observed despite the samples being processed together. This suggests that cell viability in frozen

PBMCs is probably reduced in samples by the process itself although quite at random with some being greatly affected and other less so. The extent of the effect on the quality of the flow data itself needs to be analysed separately.

Therefore, as a result of this analysis, I decided to use all samples irrespective of viable cells and to manually review all FC marker data.

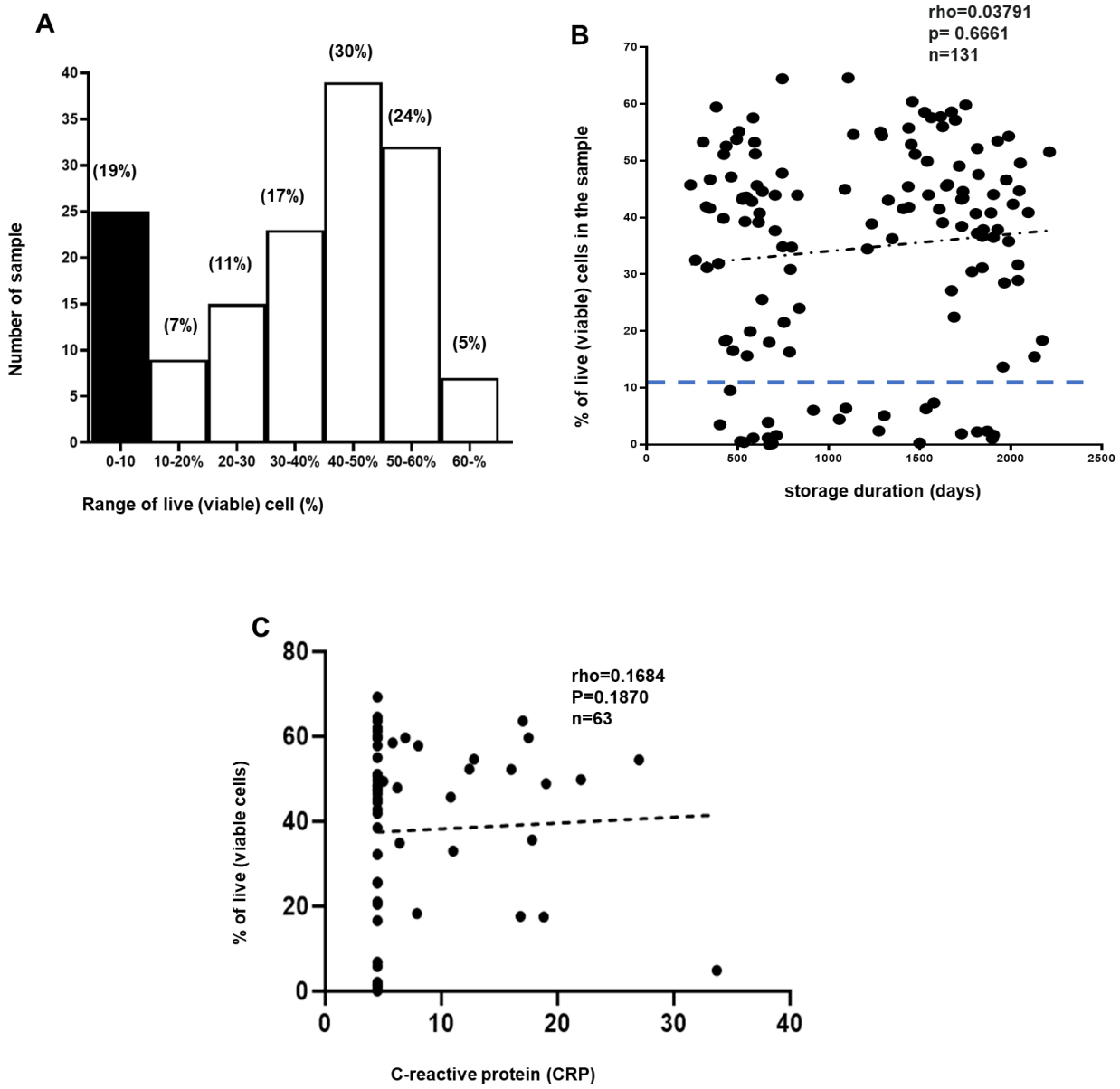


Figure 24 Cell viability in frozen sample

(A) Distribution of viable cells (%) in frozen PMBCs sample (n=131). Black bar represented samples n(%) with ≤ 10% cells viability. Plain bars indicated various proportion of sample n(%) with >10% viability.

(**B**) **Correlation between % viable cells in sample and storage duration**. Data showed no significant linear relationship between cell viability and length of storage (rho=0.03791, p= 0.6661). Blue dotted line indicating 10% cut-off point for analysable FCS data.

(**C**) Spearman correlation analysis showing no association between % live cell versus CRP (rho=0.1684, p=0.1870, n=63)

4.1.3.2 Acceptable limits in cell viability for flow cytometry analysis and loss of data using frozen PBMCs

I then proceeded to the evaluation of the effect of loss of viability on flow analysis of lymphocyte subset to determine if a loss of viability could

1. limit data quality and

2. if a cut-off should be established to determine acceptability of frozen PBMCs.

To achieve this, all samples, regardless of % cell viability contained were analysed to quantify lymphocyte lineages (CD4+T-, CD8+T-, B-, NK-, and NKT-cells), CD4+T and CD8+T-cell subsets (naïve/ IRC/ memory/Th17), B-cell subsets (Naïve/ memory/ plasmablasts/ regulatory B-cell (Breg).

Representative flow plots showing samples with varying % of cell viability analysed for all 4 panels are presented **(Figure 25 to Figure 29).**

Critically, none of the lineage lymphocyte and the lymphocyte subsets could be quantified in samples with very low levels of cell viability <5% (all 4 panels). This suggests outright rejection of these samples was needed, and this allowed me to rule out any samples with ≤ 5% viable cells for flow analyses of lymphocyte subsets.

For viability of between 5-15%, the result showed that the identification of cell population was possible for some but not all samples as well as variables according to panel analysed . In such lower viability samples, the lineage cell populations may be enumerated with a limited issue (**Figure 25** middle column), while this is impacting much more substantially the enumeration of T-cell (**Figure 26 and Figure 27 m**iddle columns) and B-cell subsets (**Figure 28** middle column)

as well as Th17 cell **(Figure 29** middle column). This suggests that in frozen samples with low cell viability, data should be reviewed by an experienced observer and may have to be considered manually for acceptance or rejection in flow analysis work.

In samples with greater than 15% viable cells, results presented showed that all panels performed well in generating clear cell populations and expression of different markers suggesting that samples with >15% cell viability remain suitable for flow analysis and may be considered an acceptable viability limit for analysis, although the number of cells available for analysis will be lower than in fresh sample as shown in **figure 23B**. This may need to be considered for naïve subpopulation.

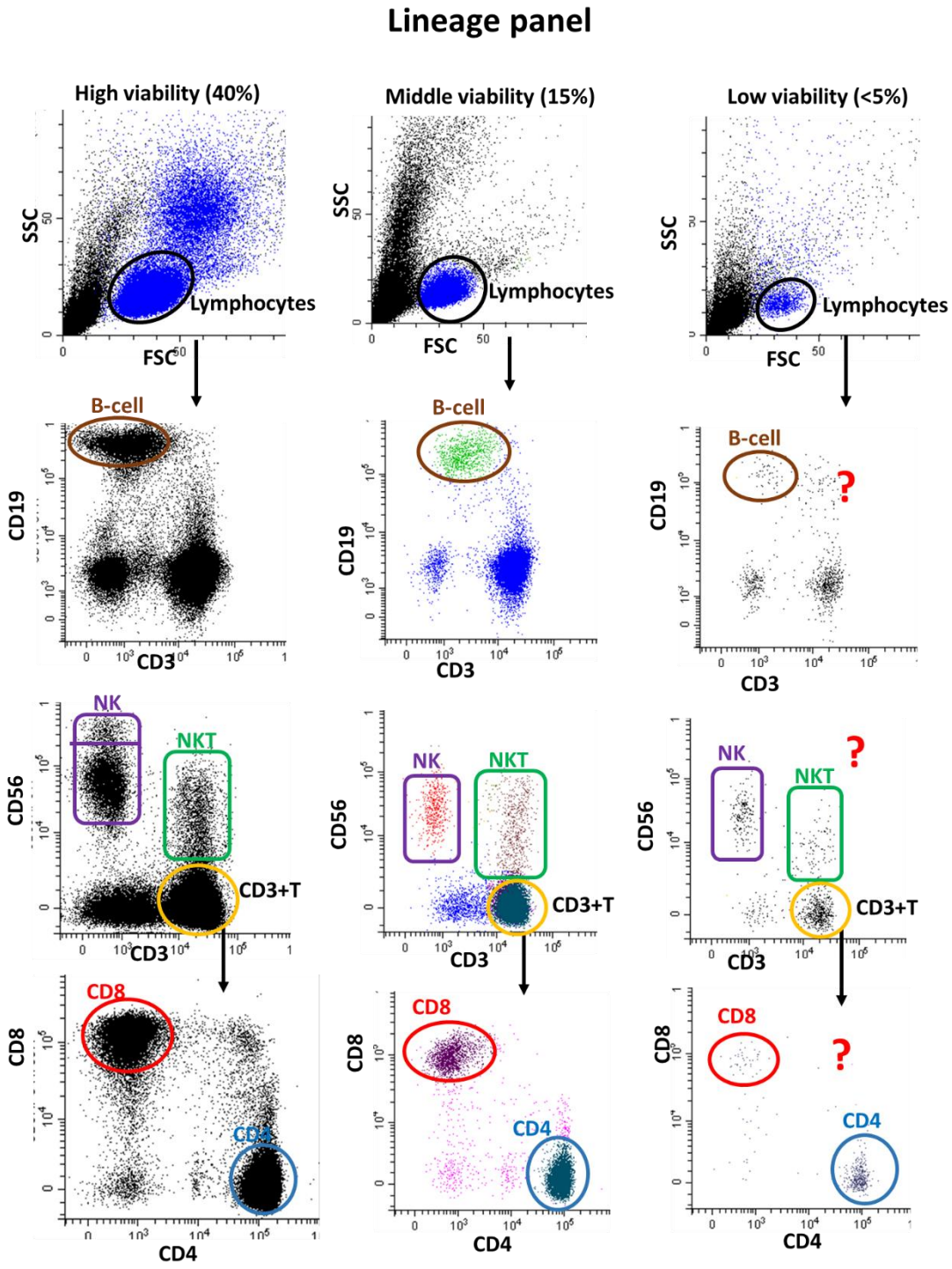


Figure 25 Representative flow plot of frozen PBMC of high, middle and low cell viability analysed for Lineage cells.

left column: Gating for PMBC with high viability (40%) showing clearly identifiable/quantifiable subsets of interest which in **Middle column** reduces for PBMC (15%) and in **right column**, completely difficult to enumerate for PMBC with < 5% viability. "?" represented plots where cells of interest were too difficult to enumerate. Gate (black circle) in FSC/SCC plot represents the lymphocyte.

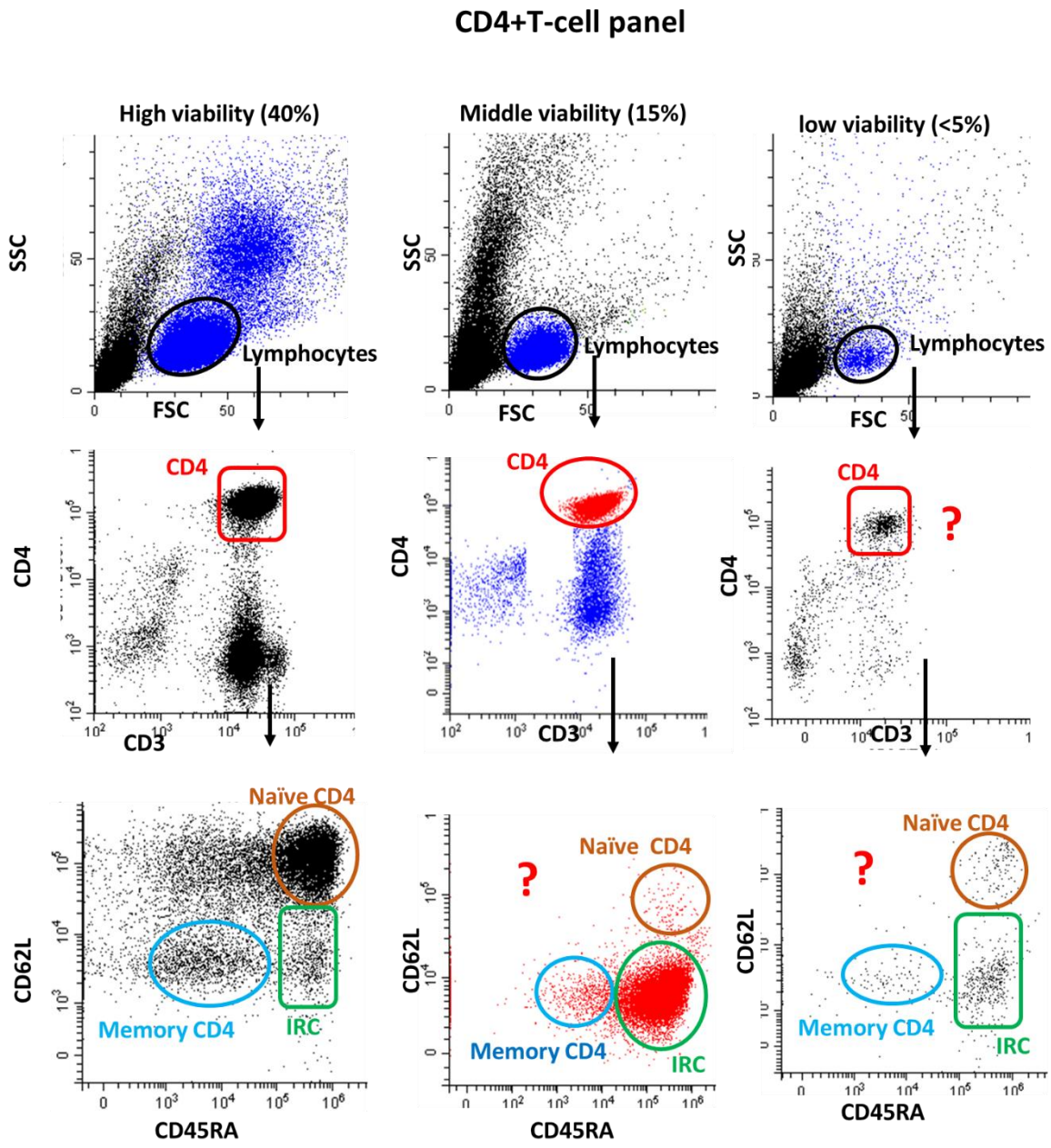


Figure 26 Representative flow plot of frozen PBMC of high, middle and low cell viability analyzed for CD4 T-cells subsets.

left column: Gating for PMBC with high viability (40%) showing clearly identifiable/quantifiable subsets of interest which in **Middle column** reduces for PBMC (15%) and in **right column**, completely difficult to enumerate for PMBC with < 5% viability. "?" represented plots where cells of interest were too difficult to enumerate. Gate (black circle) in FSC/SCC plot represents the lymphocyte.

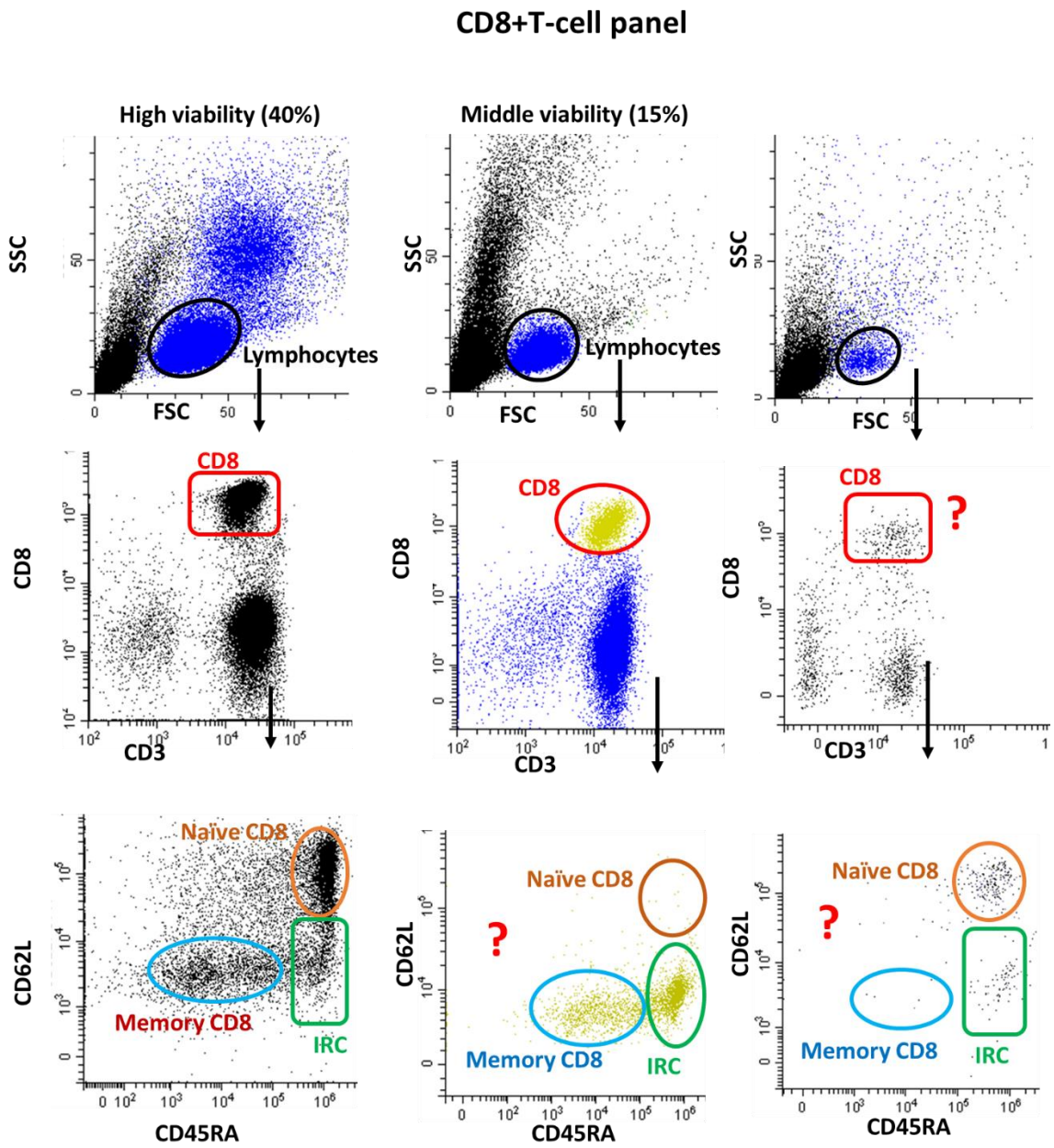


Figure 27 Representative flow plot of frozen PBMC of high, middle and low cell viability analyzed for CD8 T-cells subsets.

left column Gating for PMBC with high viability (40%) showing clearly identifiable/quantifiable subsets of interest which in **Middle column** reduces for PBMC (15%) and in **right column**, completely difficult to enumerate for PMBC with < 5% viability. "?" represented plots where cells of interest were too difficult to enumerate. Gate (black circle) in FSC/SCC plot represents the lymphocyte.

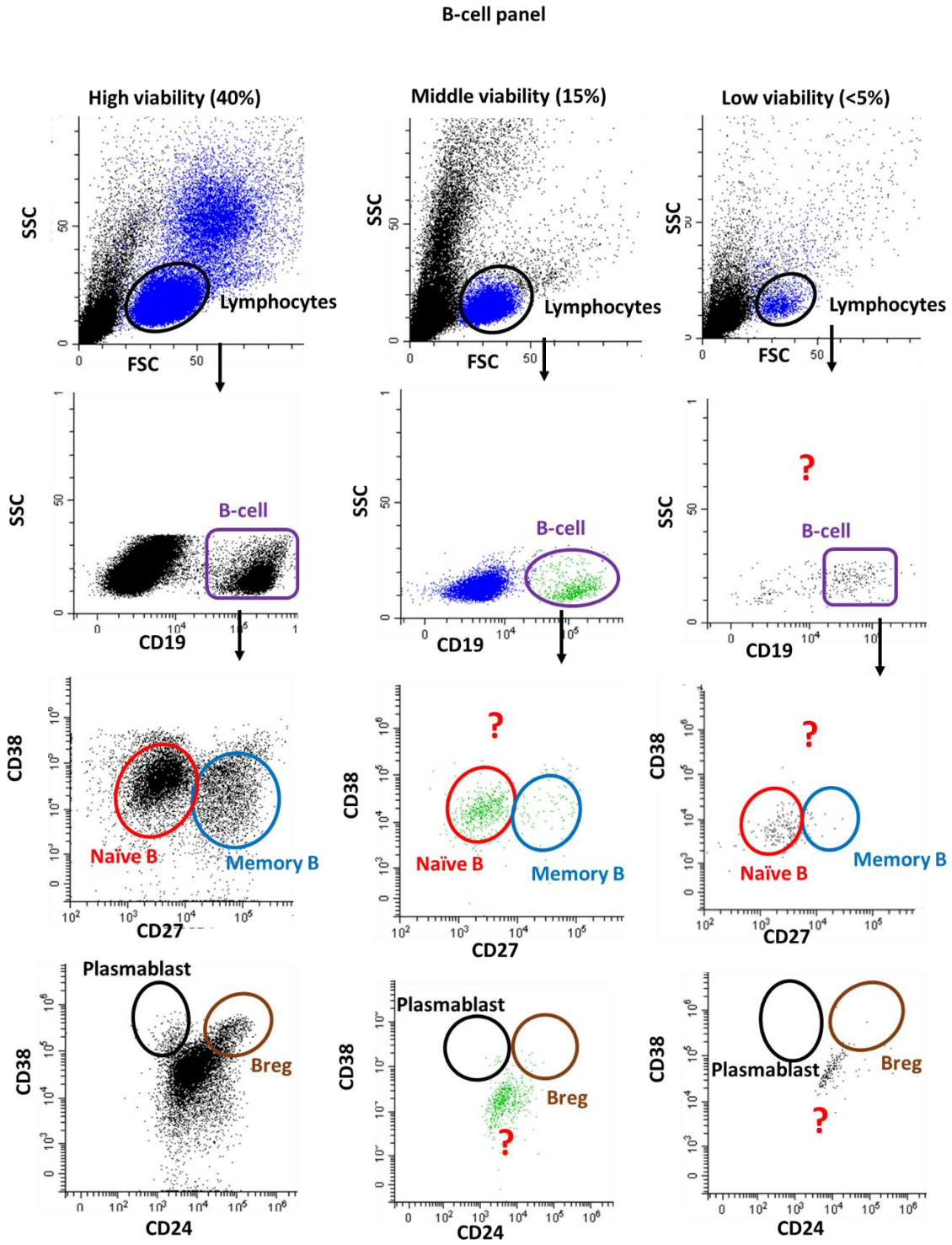


Figure 28 Representative flow plot of frozen PBMC of high, middle and low cell viability analyzed for B-cells subsets.

left column: Gating for PMBC with high viability (40%) showing identifiable/quantifiable subsets of interest which in **Middle column** reduces for PBMC (15%) and in **right column**, completely difficult to enumerate for PMBC with < 5% viability. "?" represented plots where cells of interest were too difficult to enumerate. The gate (black circle) in the FSC/SCC plot represents the lymphocyte**.**

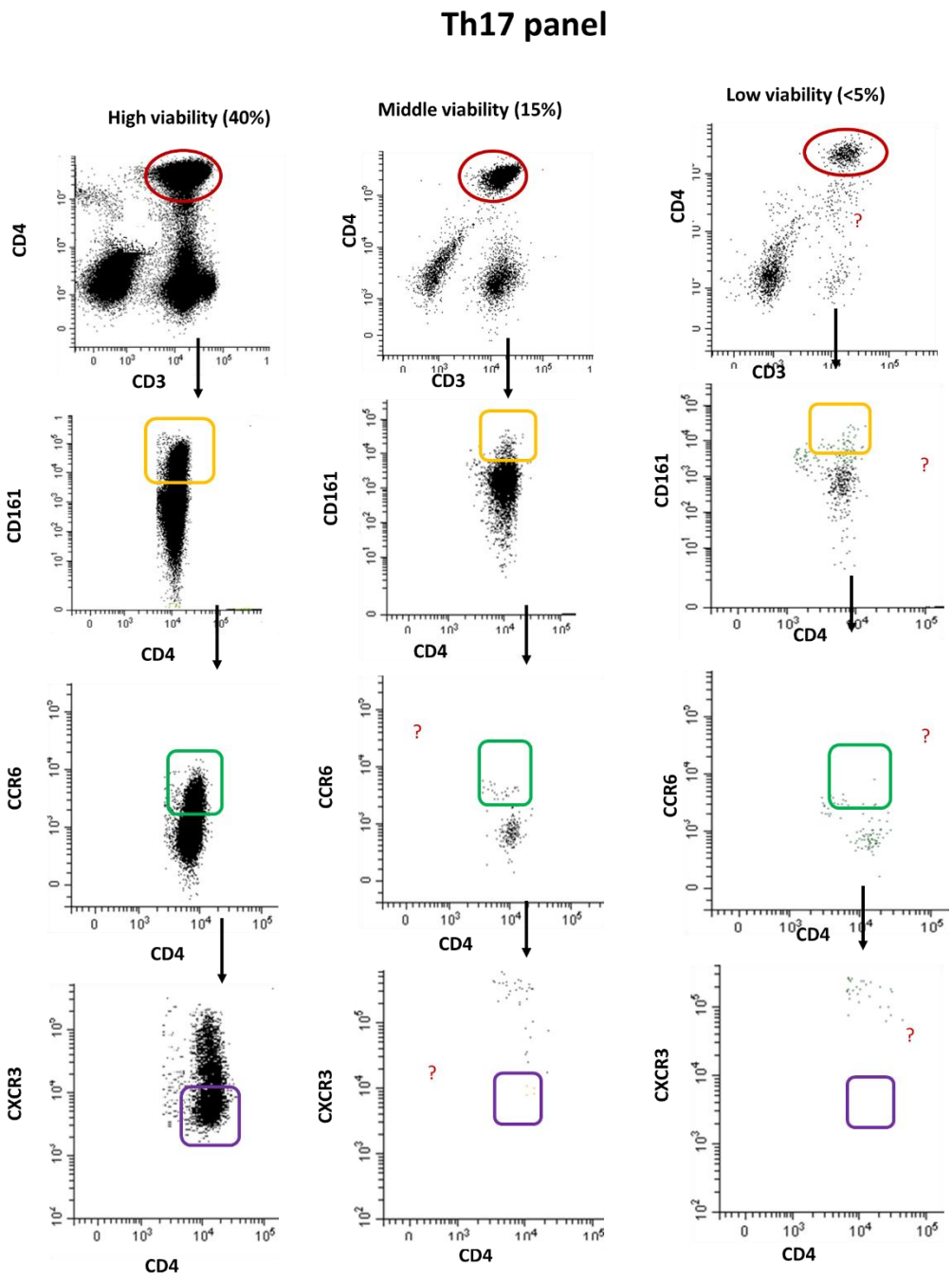


Figure 29 Representative flow plot of frozen PBMC of high, middle and low cell viability analysed for Th17 cell.

Left column: Gating for PMBC with high viability (40%) showing clearly identifiable/quantifiable subsets of interest which in **Middle column** for PBMC (15%) and **right column** (< 5% viability), were completely difficult to enumerate. "?" represented plots where cells of interest were too difficult to enumerate. Gate (black circle) in FSC/SCC plot represents the lymphocyte. CD4+T cell (gated in red), CD4+161+ cells (gated in yellow), CD4+161+CCR6+ (gated in green), and Th17 (defined as CD4+161+CCR6+ CXCR3- cells, gated in purple)

This resulted in 3 levels of cell viability with an apparent effect on the ability to analyse frozen sample; <5%, 5-15% and >15%, the latter being acceptable. Of all the frozen PBMC analysed, 13% of frozen samples were unacceptable in the GOLMePsA trial (no lineage or cell subset could be gated),, with an additional 5% of samples (with 5-15% viability) where by a manual inspection could still validate data to help rescue few more datapoint. However, only 82% of samples were acceptable as those could ensure high-quality data , despite the reduced number of live cells that can be used for the enumeration**.**

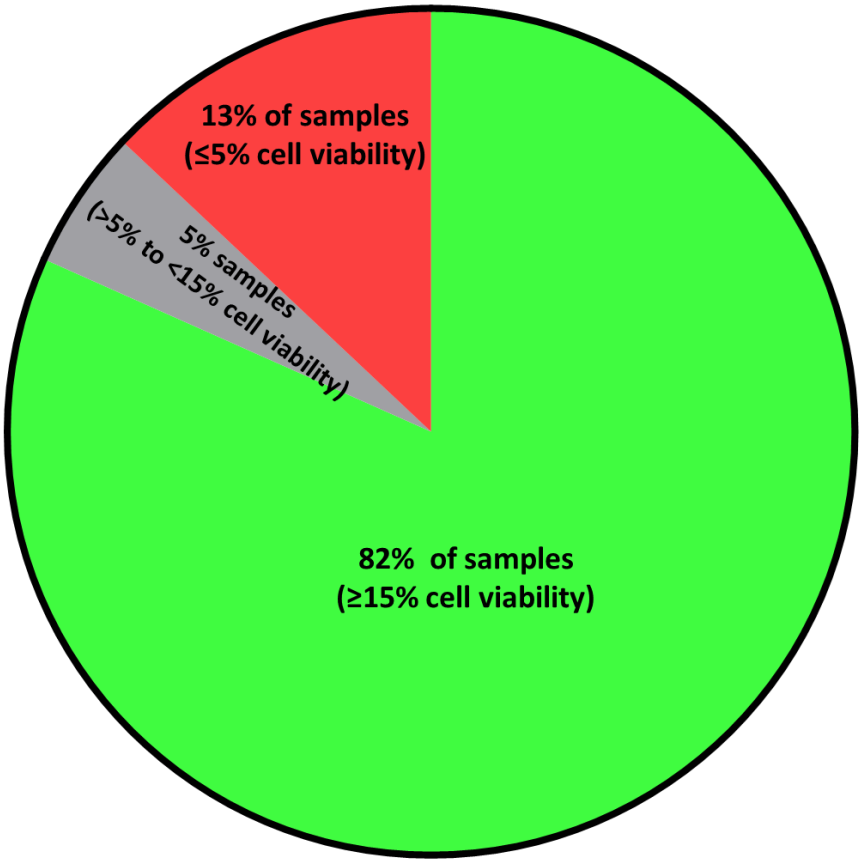


Figure 30 Acceptable % cell viability in frozen PBMC samples. Pie chart presented the three levels of acceptable % cell viability in frozen PBMC samples (n=131)). 13% of the sample (Red zone) with 5% or less viable cells indicating outright rejection as cells cannot be enumerated. 5% of sample (Gray zone) showed data with cell viability greater than 5 but less than 15% still difficult to analyse and maybe rejected as cell enumeration may not be possible for all panels. 82% of samples (Green zone) showed data with 15% or more cell viability which is perfectly suitable for flow analysis.

Overall retrospective frozen PBMCs samples showed a loss of data due to freezing in 15% of the sample. Compared to the loss of data when using fresh cells due to transport/human error/technical error (as shown in **section 4.1.1.1, page 158**). This is still allowing for the analysis of a substantial amount of data.

4.1.3.3 The effect of freezing PMBCs on flow cytometry analysis of lymphocytes and their subsets.

So far, the analysis had focused on whether the population could be clearly identified and gated but I still needed to investigate further whether the FC quantification of lymphocyte cell lineages and their subsets remains comparable to data in fresh samples. GOLMePsA being a trial of early PsA, patients were recruited from the same EAC cohort as the early RA/nonRA. Therefore, I compared the frequency of each cell and subsets from the GOLMePsA frozen samples (which had a cell viability that permitted flow analyses) with the frequency of LS obtained by analysing FSC data file previously acquired on fresh blood for GOLMePsA patients. Fresh blood for few patients were collected on the same day and some blood tubes sent by mistake to the NHS lab while other blood tubes were process as per the protocol to freeze PBMC. Match paired patients' samples (fresh versus frozen, n=15) were available and then selected for comparison analysis.

Five main cell subsets were compared between fresh and frozen sample:

- lymphocyte lineage subsets (CD4+T-, CD8+T-, B- and NK-cells),
- CD4+ T-cell subsets (naïve/IRC/memory),
- CD8+T-cell subsets (naïve/IRC/memory),
- B-cell subsets (Naïve/memory/plasmablasts/regulatory B cell(Breg),
- Th17 cells.

Impact of freezing PBMCs on lymphocytes lineage subsets quantification

Staining result fresh vs frozen PBMC

I first assessed whether the staining pattern of phenotype of lineage subsets in the frozen were comparable to the fresh sample and satisfactory for the enumeration of lineage populations in this panel. The result showed that for the LC panel (**Figure 31**) the use of frozen PBMC has relatively no impact on the detection of the B-cell , NK CD56 bright and NK CD56dim, and NKT cells and CD4+T and CD8+T cells despite using different antibody-dyes and different FC machines.

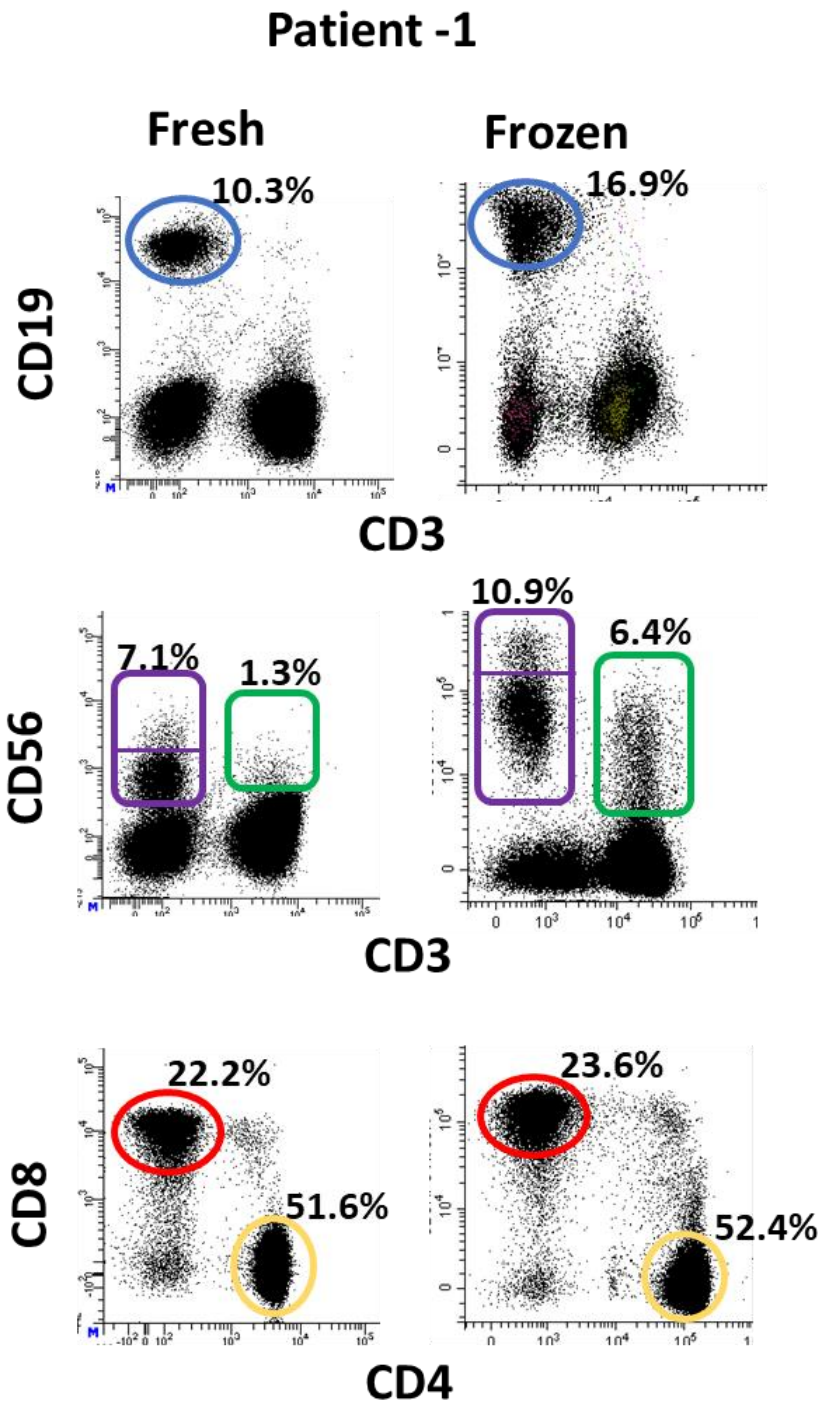


Figure 31 Representative flow cytometry plot from lineage panel comparing fresh versus frozen sample for the same patient.

Staining was performed using different antibody-dyes and data also acquired on different machines. Plots show the variation in frequencies for each cell between fresh and frozen sample. The blue gate defines the B-cells **(upper row),**, the purple gate NK cell (the low portion indicates NK CD56dim while the upper portion indicates NK CD56bright), and the green gate represents the NKT cells **(Middle row)** while red and yellow gates define CD8+T cell and CD4+T cell respectively**(last row).**

Comparison analysis results for LS (fresh vs frozen)

First, I analysed the 2 set of paired data using paired T-test and showed no significant difference for CD4 and CD8 T-cells (p=0.469 and 0.066). However, there appears to be small difference for B (p=0.003), NK(p=0.004) and NKT(p=0.001) subsets.

T-cell being the most abundant subset, the effect of freezing on the frequencies were limited while for the 3 other subsets it would be more important to verify whether these changes were randomly distributed or consistently affecting one lineage. So, I reanalysed data with match-paired data using the Wilcoxon test for the sign of difference (up versus down, **figure 32**).

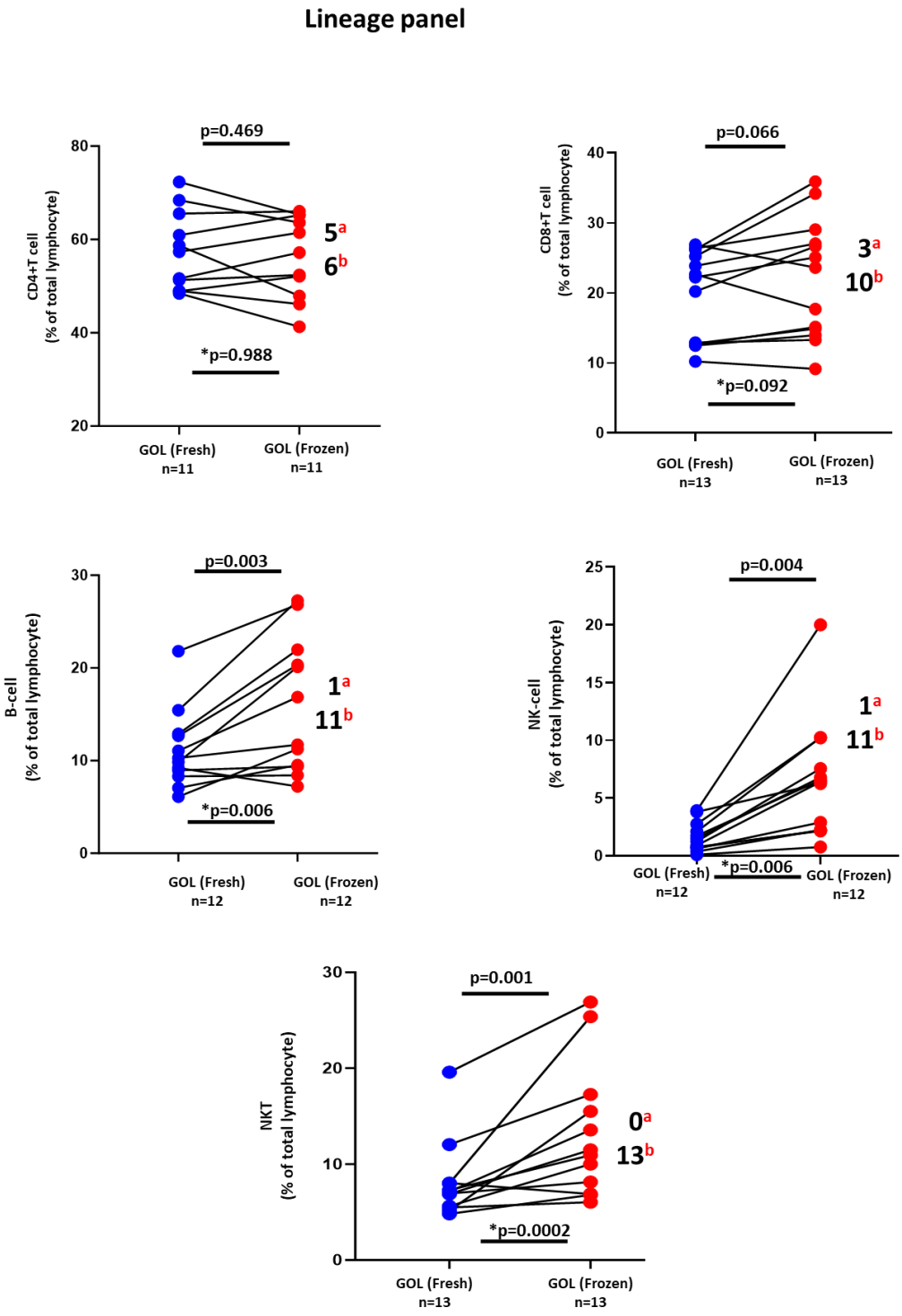


Figure 32 Comparison of fresh sample and frozen for each LS .

Fresh sample (coloured in blue) compared to frozen (Coloured in red) using Wilcoxon match-paired test for sign of the difference (up and down). CD4+T and CD8+T showed no significant difference in number of up and down pairs. For B-cells, NK cells NKT cells there were significantly more up pairs. p =T test p-value, P* = Wilcoxon test for sign of different of direction of change.

[a] frozen < Fresh(Negative Differences), [b] frozen > Fresh (Positive Differences)

The result showed no significant change in direction for CD4+T (6 up and 5 down, p = 0.988) and CD8+T (10 up and 3 down p=0.092) in frozen sample in contrast to B-cell (11 up and 1 down p = 0.006), NK-cell (11 up and 1 down p=0.006) and NKT-cell (13 up and 0 down p=0.0002) were clear higher % were often seen in frozen samples **(Figure 32).** Despite the potential impact of over 2 years of cryopreservation, our result suggests cryo-stability of CD4+T and CD8+T cells using flow cytometry while other LS tend to show higher frequency in frozen samples compared to fresh blood samples.

I then used a correlation analysis to directly compare frozen LS frequencies in fresh blood and frozen to identify where the data diverged.

The percentage of CD4+T (CD3+CD4+) showed no significant difference between frozen and fresh cells (**Figure 33**, rho =0.8455; p = 0.0018). Similarly, no difference in % were obtained for CD8+T (CD3+CD8+; rho = 0.7967, p = 0.0018).

For B-cell (CD19+CD3-, rho=0.8392, p=0.0011) results were closely correlated but with clear departure from being similar and showing consistently higher % B-cell in the frozen sample. For NKT-cells (CD56-CD3+, rho=0.8077, p=0.0014) data were also closely correlated. A non-significant correlation (Rho=0.4336, p=0.1616)) was observed for the frequency of NK-cell (CD56+ CD3-) suggesting that freezing was particularly affecting this subset enumeration.

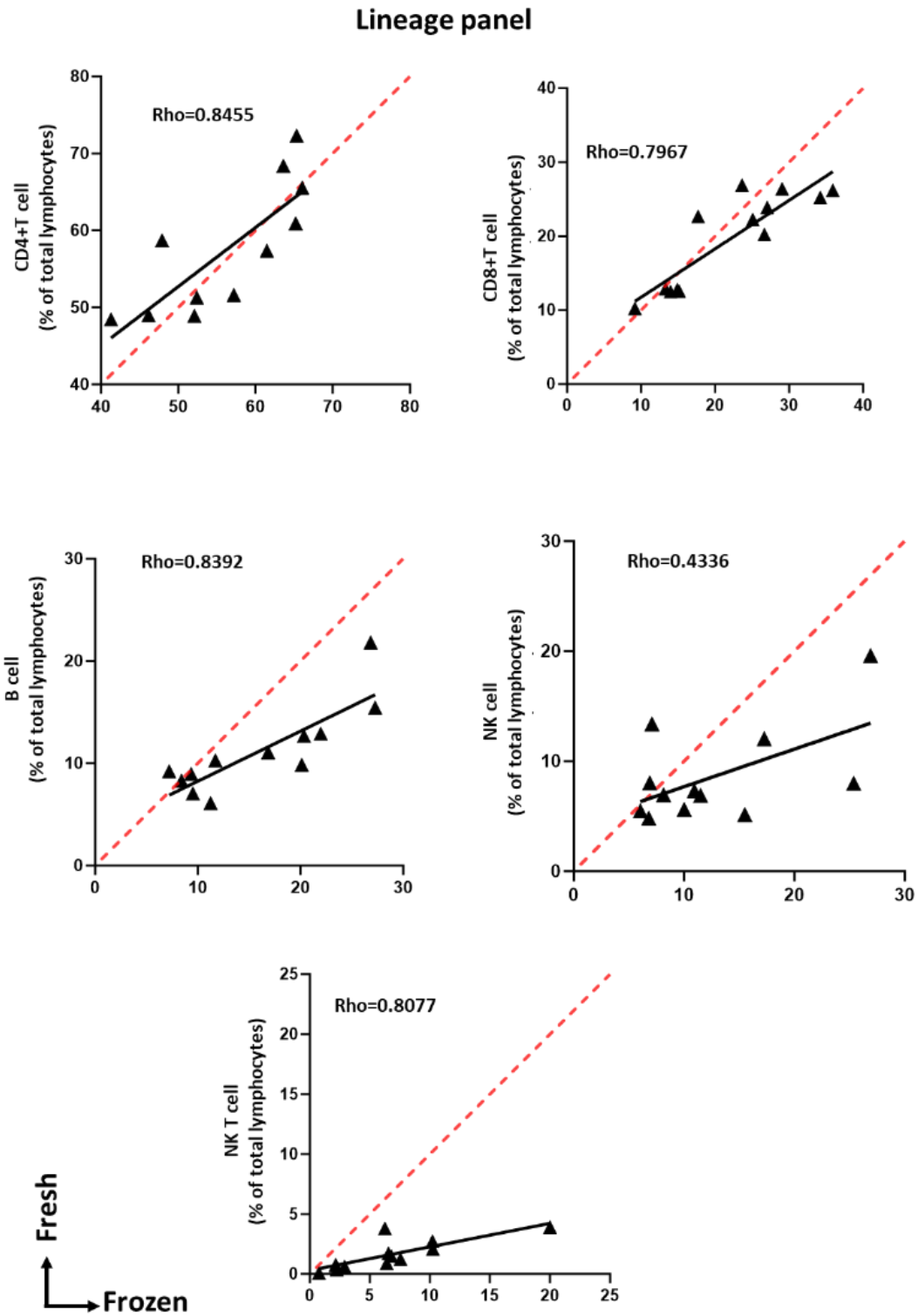


Figure 33 Correlation analysis of lineage frequencies in fresh and frozen. Frequencies of LS subsets in fresh and frozen PBMC samples from the same participants are shown in scatter plot. Complete identity line (Slope=1) between fresh and frozen cells is indicated by red dotted lines. The linear regression line of best-fit of data is indicated by black lines. Spearman correlation coefficients (rho) shown for each panel.

4.1.3.4 Altered T-cell subsets quantification in frozen samples

Staining pattern fresh vs frozen PBMC

The FC analysis for the CD4+T-cell subsets (naïve, IRC, and memory) was performed using the previously used 2-markers gating strategy (CD45RA/CD62L). I first assessed whether the staining pattern in the frozen samples was satisfactory for the enumeration of subpopulations, not compromising the separation of any T-cell subsets compared to the fresh sample. As shown in **figure 34** staining pattern was not identical, particularly for the expression of CD62L defining naïve/IRC in both the CD4+ and CD8+ subsets. Use of a frozen PBMC has a substantial impact on the detection of the naïve and IRC cell subsets. Therefore, enriching CD62L- cells.

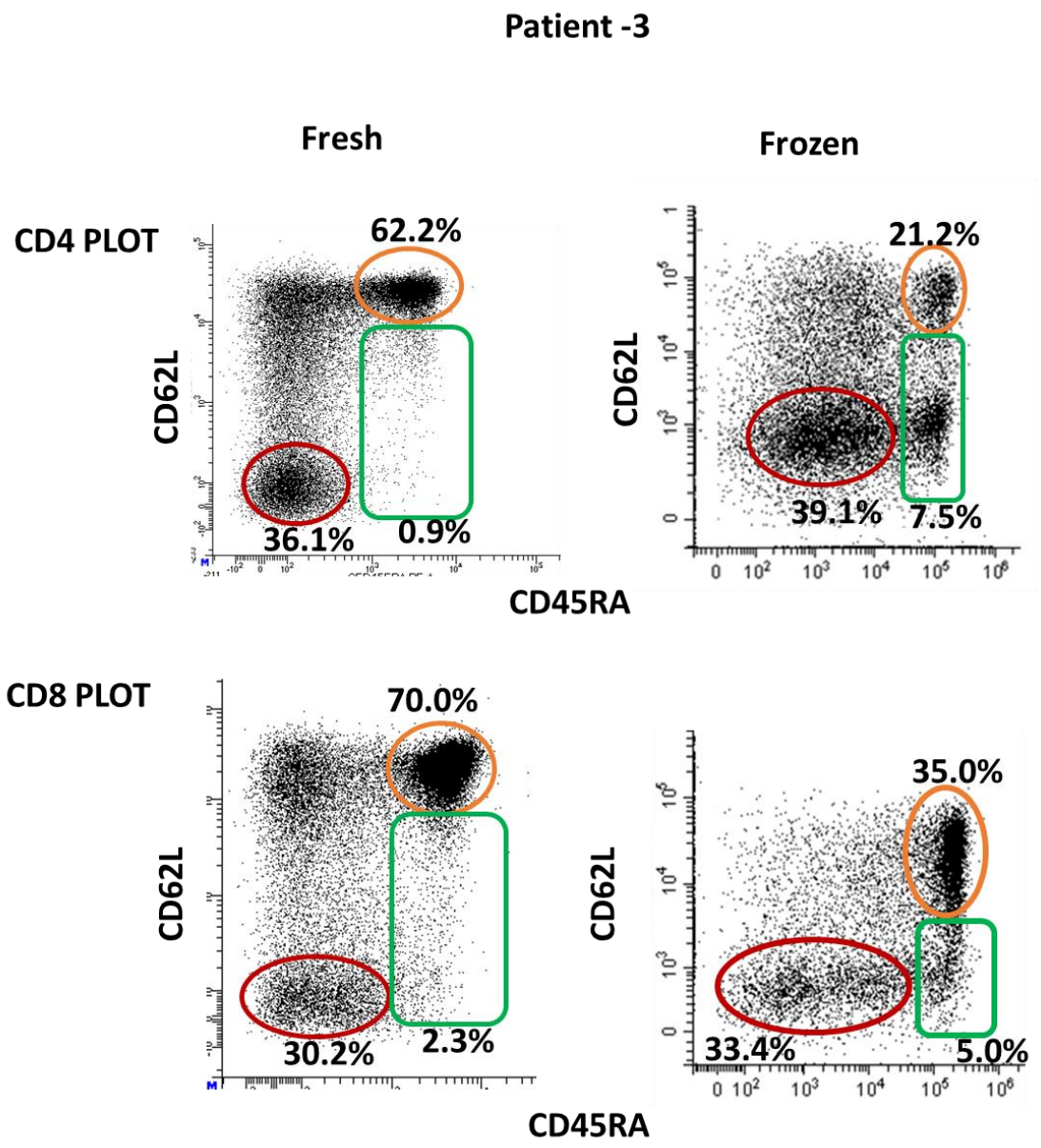


Figure 34 Representative flow cytometry plot of CD4 and CD8 T-cell panel comparing fresh versus frozen sample for the same patient.

Staining was performed using different antibody-dyes and data also acquired on different machines. Plots show the variation in frequencies for each cell subsets between fresh and frozen sample. The yellow gate defines the naïve subsets, the green gate defines the IRC, and the red gate represents the memory subset . Of note the Diva-NHS acquired data used an analogue FC scale (on a log scale) while the CytoFlex used a digital one

Comparison analysis for CD4+ T-cell subsets (fresh vs frozen)

A similar analysis was performed for the 3 CD4+T cell subsets- naïve CD4+T cell defined as CD45RA+CD62L+, IRC (CD45RA+ CD62L-), and the memory CD4+T cell (CD45RA- CD62L-).

Matched paired analysis (n = 15) performed using similar approach confirmed that naïve CD4+T in the frozen sample were significantly different from fresh samples (T-test, p=0.0002) and this change was due to reduced naïve cell frequency (Wilcoxon test, 0 up and 15 down, p=0.0007, **figure 35**). Following the reduction in naïve, the IRC subset was found to be significantly different (T-test, p=0.0005) and increased in frozen compared to fresh samples (Wilcoxon test,14 up and 0 down, p=0.0009). This is showing a freezing-specific alteration associated with the naïve/IRC subsets possibly doe to loss of CD62L. For memory CD4+T, results confirmed the different (T-test, p<0.0001) and a significant increased frequencies (Wilcoxon test, 15 up and 0 down, p=0.0007, in frozen samples **(Figure 35).**

In addition loss of consistency, particularly poor correlations were also observed for both naïve CD4+T (rho=0.3750, p=0.1692) and IRC CD4+T (rho=0.0484, p=0.8702) between fresh and frozen samples, in contrast to memory CD4+T where a better correlation between frozen and fresh samples (rho=0.6036, p=0.0195 (**Figure 35** right hand side) was seen suggesting that memory subset frequency in the frozen sample may relatively be more stable.

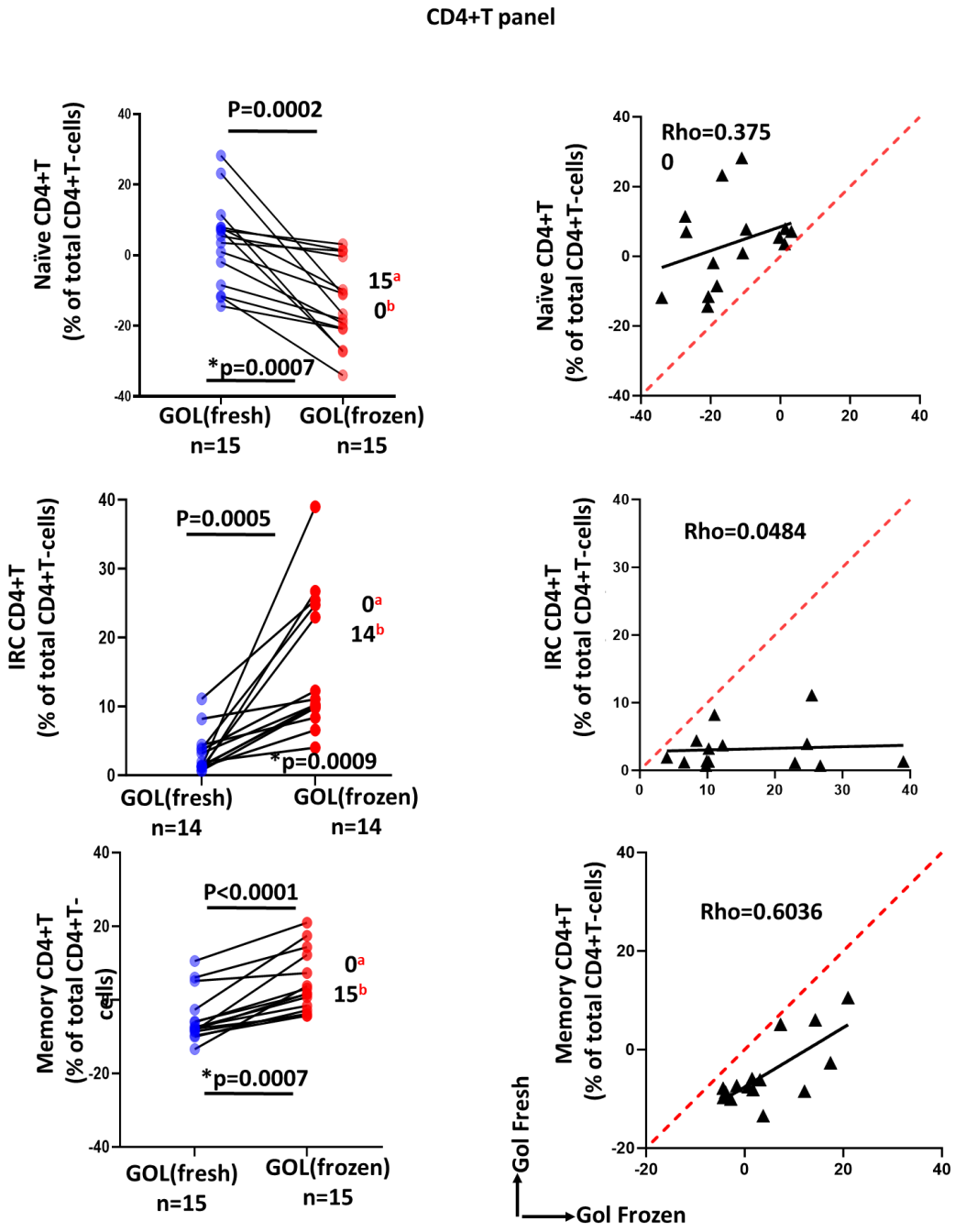


Figure 35 Comparison fresh and frozen samples for CD4+T cell subsets (based on conventional 2 markers gating (CD45RA/CD62L).

(**right column**) Spearman correlation analysis: Frequencies of CD4+T cell subsets in fresh and frozen PBMC samples from the same participants are shown in scatter plot. Complete identity line (Slope=1) indicated by red dotted lines. Linear regression line of best-fit of data is indicated by black lines. Spearman correlation coefficients(rho) are shown for each subset. (**Left column**) Fresh sample (coloured in blue) compared to frozen (Coloured in red) using Wilcoxon match-paired test for sign of the difference (up and down). Naïve showed significant difference in number down pairs while IRC showed more up pairs. For memory there were significantly more up pairs. p=T test p-value, P* = Wilcoxon test for sign of different of direction of change. [a] frozen < Fresh (Negative Differences), [b] frozen > Fresh (Positive Differences)

CD4+T cell subsets (rescuing the analysis)

Changes in CD62L surface marker expression with freezing have been reported with specific shedding from the cell surface being the main explanation [723, 724]. CD62L pattern of expression was indeed different between fresh and frozen samples **(Figure 34, page 202),** and the frequency changes are mainly due to the loss of CD62L expression in frozen samples

Considering that IRC is present in the CD4+ T cells subset but in a small proportion (except in a few cases with very high inflammation), I proposed a 1-marker identification gate for enumerating naïve CD4+T (CD45RA+) in frozen samples as a potential solution to help fix the issue observed with using the 2-marker phenotype (CD45RA+CD62L+). I therefore quantified again naïve CD4+T using only the CD45RA+ marker. Both subsets were then corrected for age gating this new CD45+ gate in consideration.

The 1-marker identification strategy was applied further to the fresh, and frozen samples and the data compared again (**Figure 36**, n=14). Results showed that naïve CD4+T for the paired fresh-frozen matched sample now compared well (T-test, p=0.057) with significant correlation value (rho=0.8330, p=0.0004) and no significant direction of change (Wilcoxon test, 5 up and 9 down p=0.424). The loss/change of CD62L expression due to freezing was therefore leading to the negative skewness in naïve CD4+T enumeration using the conventional 2 markers (CD45RA/CD62L) and was overcome using the 1-marker CD45RA+ gate, although at the expense of being able to also quantify IRCs.

For memory CD4+Tcells (CD45RA- cells), the 1-marker gating strategy also showed no difference (T-test=0.058), maintained a good correlation (rho=0.7319, p=0.0040) (**Figure 36**) and confirmed no significant direction of change (Wilcoxon test, 9 up and 5 down, p=0.496) between fresh sample and frozen PBMC. This suggests that memory CD4+T enumeration using the 1-marker strategy may provide a solution to fix the impact of freezing on this subset. However, although with a slight widening of data distribution in frozen samples.

Therefore, the use of a 1-marker strategy removing CD62L to enumerate naïve and memory CD4+T in the frozen sample is likely to result in data that was relatively similar.

The CD45RA (+ and -) subsets are not equivalent to the 2-markers subset (CD62L/CD45RA) used in previous analysis and data interpretation will have to take this into consideration. The shedding of CD62L due to freezing remains an issue if IRC is to be measured.

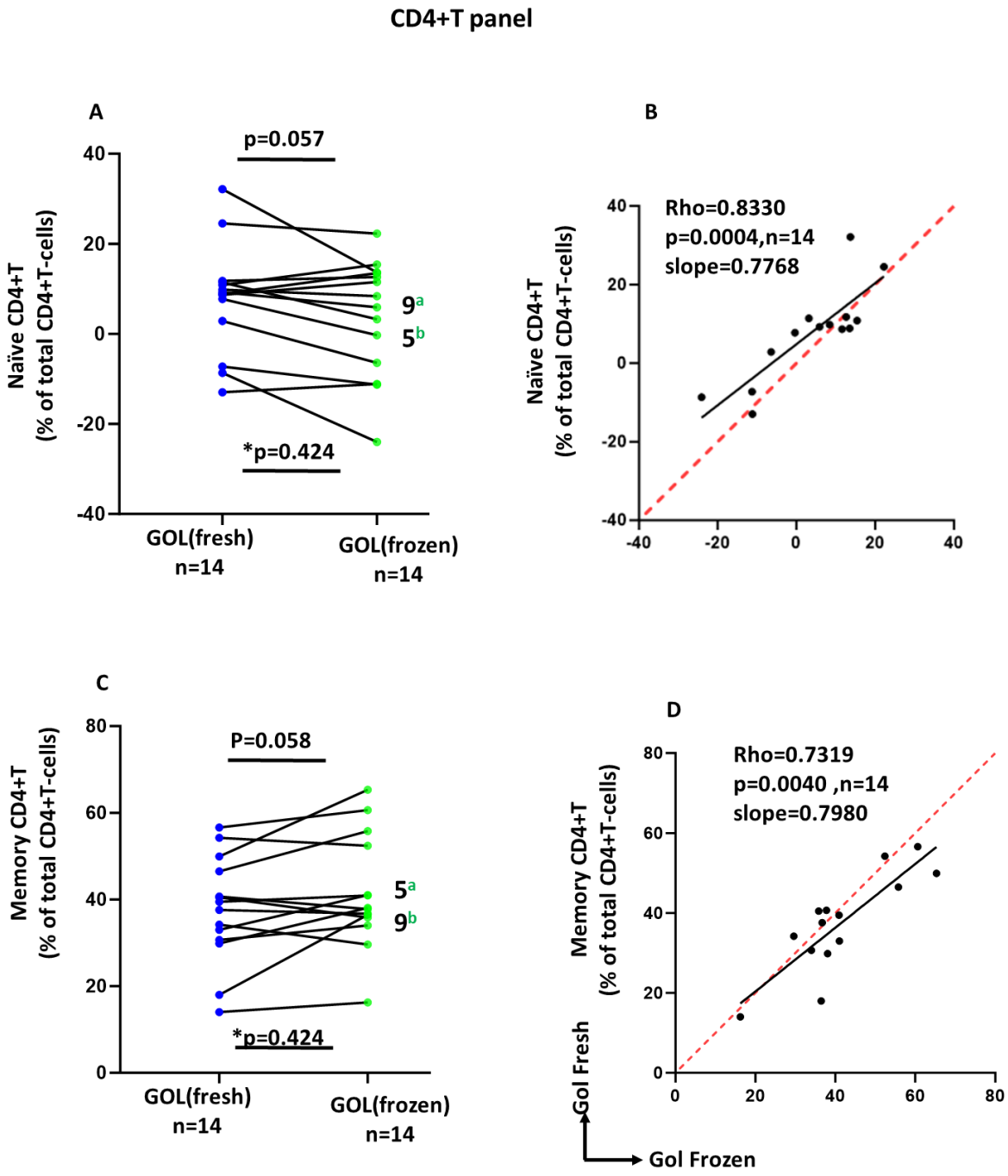


Figure 36 Comparison of CD4+T cell subset in fresh and frozen samples (based on proposed 1-marker (CD45RA).

Right column) Spearman correlation analysis: Frequencies of CD4+T cell subsets in fresh and frozen PBMC samples from the same participants are shown in scatter plot. Complete identity line (Slope=1) indicated by red dotted lines. The linear regression line of best-fit of data is indicated by black lines. Spearman correlation coefficients(rho) are shown for each subset. **(Left column**) Fresh sample (coloured in blue) compared to frozen (Coloured in red) using Wilcoxon match-paired test for sign of the difference (up and down). Naïve/memory showed no significant difference in number down pairs. p=T test p-value, P* = Wilcoxon test for sign of different of direction of change. [a] frozen < Fresh(Negative Differences), [b] frozen > Fresh (Positive Differences)

Comparison analysis for CD8+ T-cell subsets (fresh vs frozen)

CD8+T cell subsets were impacted in the same way as CD4+T-cell subsets by the shedding of CD62L. Naive CD8+T showed a significant difference (T-test, p=0.015), no significant correlation between frozen samples and fresh samples (rho=0,34851, p=0.2021) but a significant decreased direction of change in frozen sample (Wilcoxon test, 2 up and 13 down, p=0.007). This was mirrored in increased IRC CD8+T **(Figure 37**, t-test, p<0.0001, and rho=0.1226, p=0.7037) suggesting a substantial alteration in the frequency of these subsets in frozen PBMC. Again, for the memory subset, frozen samples showed a better correlation with fresh samples (rho=0.6364, p=0.0299, **figure 37**) and no significant difference (T-test, p=0.341) and no direction of change (Wilcoxon test , 6 up and 6 down, p=0.988).

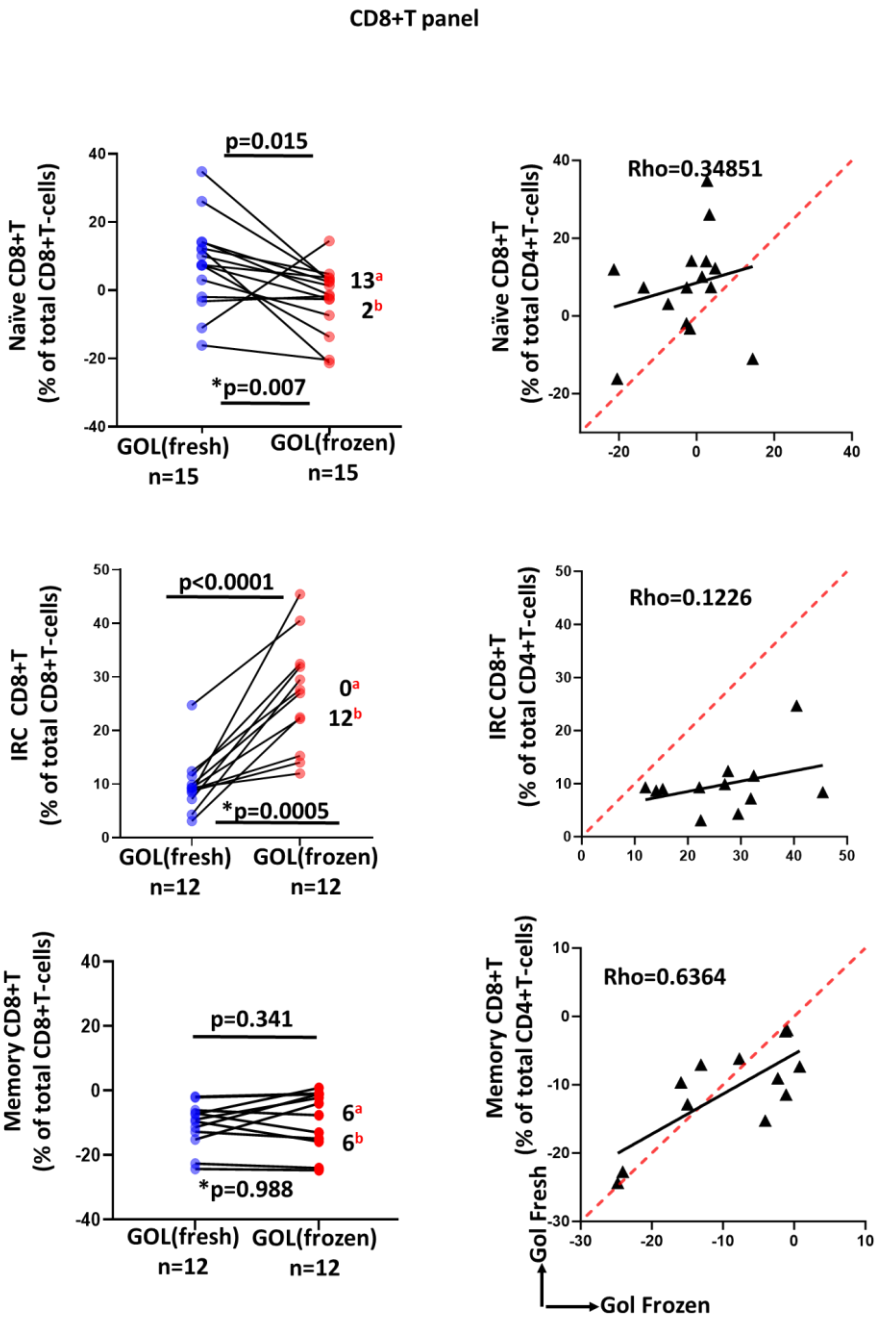


Figure 37 Comparison fresh and frozen samples for CD8+T cell subsets (based on conventional 2 markers gating (CD45RA/CD62L).

(**right column**) Spearman correlation analysis: Frequencies of CD8+T cell subsets in fresh and frozen PBMC samples from the same participants are shown in scatter plot. Complete identity line (Slope=1) indicated by red dotted lines. Linear regression line of best-fit of data is indicated by black lines. Spearman correlation coefficients(rho) are shown for each subset. (**Left column**) Fresh sample (coloured in blue) compared to frozen (Coloured in red) using Wilcoxon match-paired test for sign of the difference (up and down). Naïve showed significant difference in number down pairs while IRC showed more up pairs. For memory there was no significant difference for up and down pairs. p=T test p-value, P* = Wilcoxon test for sign of different of direction of change. [a] frozen < Fresh(Negative Differences), [b] frozen > Fresh (Positive Differences)

CD8+T cell subsets (rescuing the analysis)

Next, I repeated the 1-marker strategy (CD45RA+) to determine if it could also find a similar solution for CD8+T cell subsets (

Figure 38). For naïve cells, the 1-marker strategy did not compare well in paired fresh-frozen matched sample (T-test, p=0.035), with a poor correlation result (rho= 0.0945, p=0.6136, n=14) and no significant direction of change for naïve CD8+T cells (Wilcoxon test, 9 up and 5 down, p=0.424). However, the overall significant increase for naïve cell in frozen sample (T-test, p=0.035) indicates that the use of 1-marker strategy to enumerate naïve CD8+T in frozen sample is likely to result in over-enumeration, suggesting this is unlikely to offer a solution to fix the impact of freezing on this subset. The CD45RA+ cells tend to contain a much higher frequency of IRC in CD8+T-cells (notably compared to CD4+T-cells) and in fresh samples as such this is more likely to affect the results. Data interpretation will have to take this into high consideration.

For memory CD8+T, the 1-marker strategy was not used as no difference was found in frequency between fresh and frozen samples.

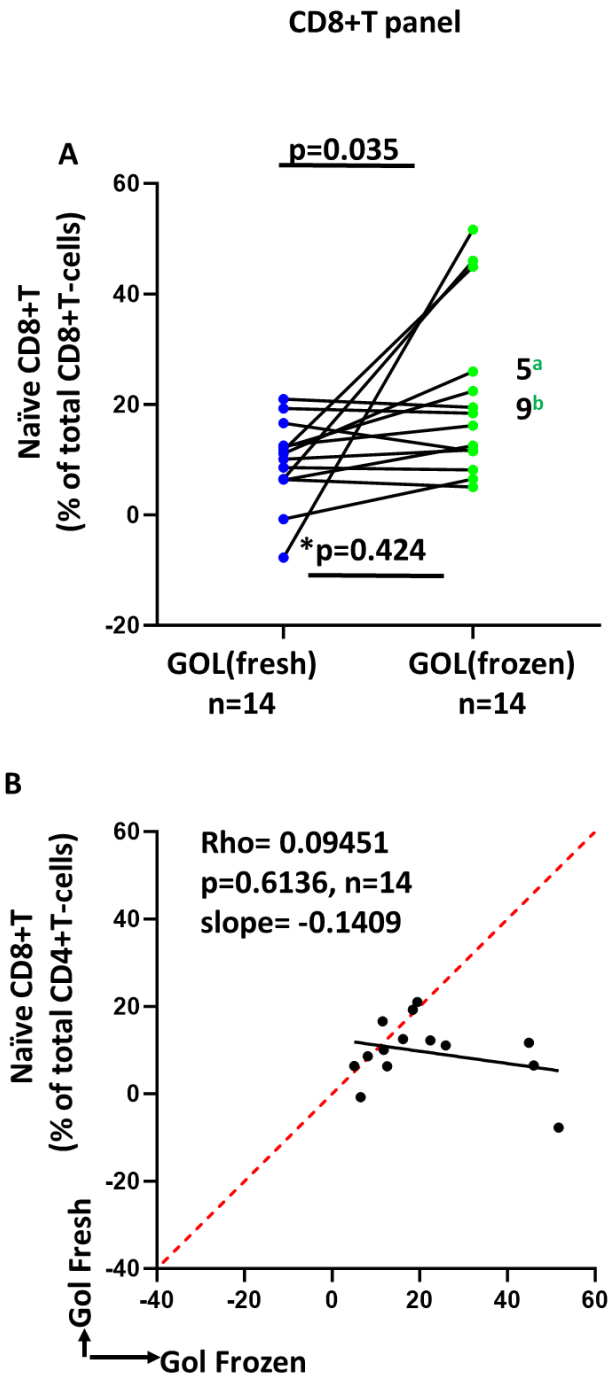


Figure 38 Comparison of CD8+T cell subset in fresh and frozen samples (based on proposed 1-marker (CD45RA).

(A) Spearman correlation analysis: Frequencies of naiveCD4+T cell subset (fresh versus frozen) from the same participants are shown in scatter plot. Complete identity line (Slope=1) indicated by red dotted lines. The linear regression line of best-fit of data is indicated by black lines. Spearman correlation coefficients(rho) are shown. (B) Fresh sample (coloured in blue) compared to frozen (Coloured in red) using Wilcoxon match-paired test for sign of the difference (up and down). Naïve showed a significant difference in number up pairs. p=T test p-value, P* = Wilcoxon test for sign of different of direction of change. [a] frozen < Fresh(Negative Differences), [b] frozen > Fresh (Positive Differences)

4.1.3.5 Stability of B cell subsets quantification in frozen versus fresh samples

Staining pattern fresh versus frozen

For the B-cell panel, the staining pattern was also assessed to see if it was satisfactory for the enumeration of subsets. The result presented in **figure 39** showed that for B-cells, the use of a frozen PBMC has limited impact on the detection of the naïve, memory, Breg, and plasmablasts subsets, although the pattern was not fully identical, as the different FC machine/Antibodies clone/dyes also had an impact on the technique.

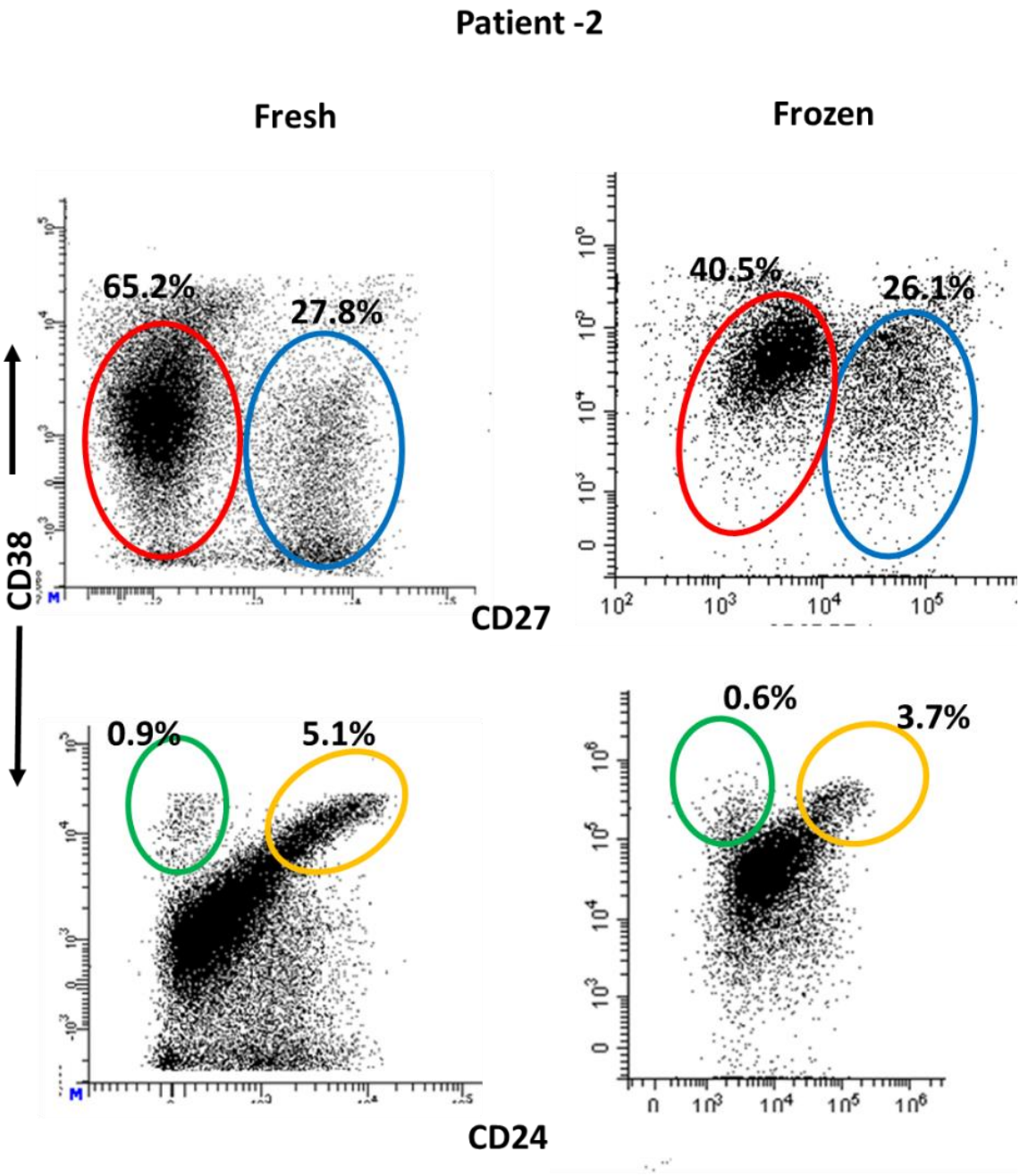


Figure 39 Representative flow cytometry plot of B-cell panel for the fresh versus frozen sample of the same patient.

Plots show the variation in frequencies for each cell between fresh and frozen sample. The red gates define the naïve B-cells , the blue gates define the memory B (upper row). The green gates represent the plasmablasts while yellow gates define Breg respectively.

Comparison analysis for B-cell subsets (fresh vs frozen)

Naive B-cell showed significant difference (T-test, p=0.015), a significant correlation between frozen samples and fresh samples (rho=0. 6845, p=0.002), and no significant direction of change for lower frequency for naïve B-cell in frozen sample (Wilcoxon test, 5 up and 10 down, p=0.302, **figure 40**).

In contrast, there were no significant difference for memory B-cell (t-test, p=0.677), Breg (t-test, p=0.329), and plasmablasts (t-test, p=0.224), and a significant correlation observed for memory (rho=0.6615), Breg (rho=0.7026), and plasmablasts (rho=0.5481) between frozen and fresh sample and no direct of change for memory (p=0.791), Breg (p=0.424), and plasmablasts (p=0.968, **figure 40**. The randomness observed in change of direction subsets % suggests a non-specific freezing impact on the B-cell subsets. Therefore, freezing may unlikely impact the frequencies of the B-cell subsets substantially notably the enumeration of memory, Breg and plasmablasts.

B-cell panel

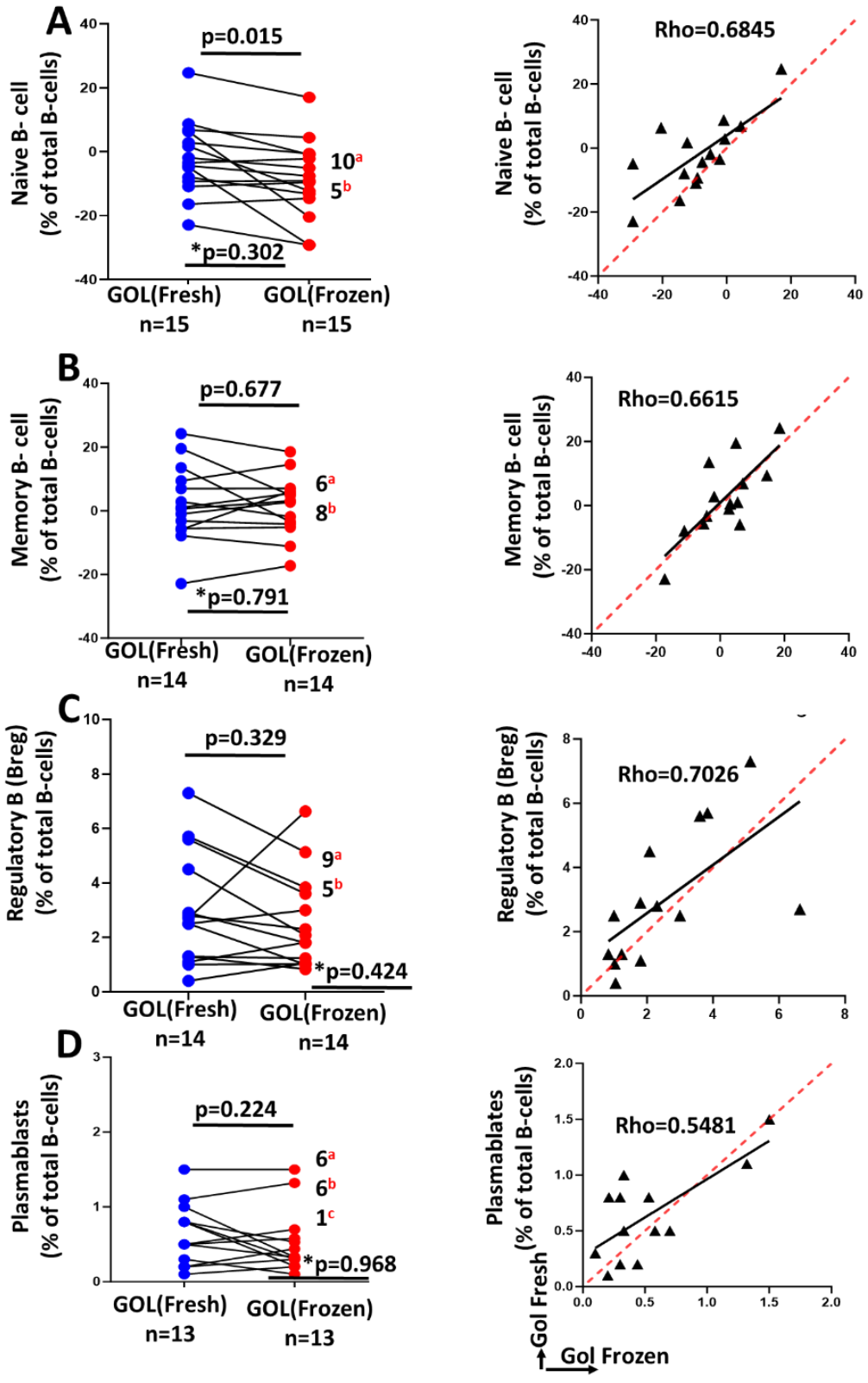

A
Naive B- cell (% of total B-cells)
p=0.015
10a
5b
*p=0.302
GOL(Fresh) n=15
GOL(Frozen) n=15
Rho=0.6845
B
Memory B- cell (% of total B-cells)
p=0.677
6a
8b
*p=0.791
GOL(Fresh) n=14
GOL(Frozen) n=14
Rho=0.6615
C
Regulatory B (Breg) (% of total B-cells)
p=0.329
9a
5b
*p=0.424
GOL(Fresh) n=14
GOL(Frozen) n=14
Rho=0.7026
D
Plasmablasts (% of total B-cells)
p=0.224
6a
6b
1c
*p=0.968
GOL(Fresh) n=13
GOL(Frozen) n=13
Plasmablates (% of total B-cells)
Rho=0.5481
Gol Fresh
Gol Frozen

Figure 40 Correlation analysis of B-cell subsets in fresh and frozen samples from GoI trial.

(Right column) Frequencies of B-cell subsets in fresh and frozen PBMC samples from the same participants are shown in scatter plot. Complete identity line (Slope=1) between fresh and frozen cells is indicated by red dotted lines. Linear regression line of best-fit of data is indicated by black lines. Spearman correlation coefficients(rho)are shown for each subset.**(Left column) comparison of B-cell subsets in fresh and frozen sample**. Fresh sample (coloured in blue) compared to frozen (Coloured in red) using Wilcoxon match-paired test for sign of the difference (up and down). Naïve showed a significant difference in number down pairs whereas other subsets did not show any significant difference for up or down pairs. [a] frozen < Fresh(Negative Differences), [b] frozen > Fresh (Positive Differences)

4.1.3.6 Impact of freezing PBMCs on analysis of Th17 cells

There was no available paired fresh samples to assess the impact of freezing on Th17 staining pattern. As such, the result of representative flow plots for assessing staining patterns could not be displayed. However, I obtained results from a Th17 panel from late PsA (n=24) from fresh samples (no data available in early PsA) from an unpublished study. The Th17 subsets showed significantly lower frequencies in frozen GOL samples compared to fresh PsA ($p<0.001$, n=24, **figure 41**), suggesting that th17 range was not closely comparable between frozen and fresh samples as could be expected. The use of late PsA may also suggest a disease stage-specific change as an alternative explanation to a freezing-specific difference. This prevented me from validating the use of frozen samples for the Th17 panel and therefore suggests that we will have to take data with this limitation into consideration when relating it to biological findings and/or trial outcomes. In addition, this may also suggest that discrepancies in the literature may be related to the use of frozen versus fresh samples

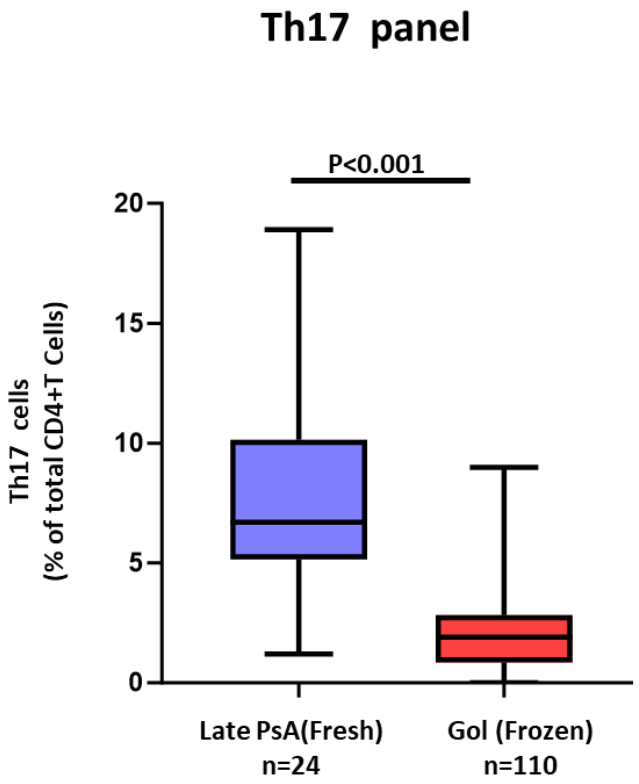


Figure 41 Changes in Th17 cells in fresh versus frozen disease samples.

Data presented in boxplot showed fresh non-RA sample (blue, n=24) compared to frozen GoI samples (red, n=110) (MWU test, p<0.001)

4.2 Clinical utility of biomarker using the FC technology

In this section of my PhD, I plan to validate the use of FC in CD4+ T-cells as well as other LS as biomarkers of disease progression across different stages of the IAC. The first part of this project (at-risk cohort) has been published [725]. I performed most of the data analysis for FC and statistical modelling while my coauthors provided a clinical overview and contributed to the manuscript editing and revision. The 2nd part (EAC cohort) and 3rd part (MXT-treated RA cohort) form the basis of data to be included in a new publication plan for 2024. For this 2nd paper, I performed most of the data analysis for FC and statistical modelling while my coauthors provided small cohort data analysis (master students), and clinical data from the cohort (ESREP students).

4.2.1 Clinical utility of (fresh blood) flow cytometry biomarker across the IAC

4.2.1.1 The phases of the IAC; cohort description

The involvement of immune cells in RA pathogenesis has been extensively described (T/B/NK-cells, monocytes [583, 726-729]. The specific cellular and molecular events that influence progression to the next stage of the IAC remain unclear [730].

Rheumatoid arthritis (RA) is a chronic autoimmune, inflammatory joint disease. A pre-clinical phase of RA has been identified, also known as the at-risk phase of the inflammatory arthritis continuum (IAC) [308, 309]. The at-risk phase can last up to 15 years, during which genetic and environmental factors contribute to the progression including a break in tolerance and the development of systemic autoimmunity manifested by the presence of autoantibodies (particularly, anti-citrullinated peptide antibodies-ACPA).

The development of pain and other musculoskeletal symptoms in the absence of synovitis precedes a final stage when synovitis develops. Treatment is conventionally initiated upon the detection of clinical synovitis.

As previously established by our group in ACPA+ at-risk individuals, three CD4+T-cell subsets (naïve, IRC-CD4, and Treg) had good predictive value as biomarkers for progression to IA, both individually and when combined with clinical variables [3, 5, 362, 520]. They were also able to predict the progression towards RA from an early arthritis clinic and to predict MXT-induced remission in early RA [5, 362, 520]. However, the predictive value of other immune cells across the 3 stages of the IAC is yet to be fully explored.

Here, I tested the hypothesis that the dysregulation of other LS, in addition to the three CD4+T-cell subsets previously reported, may provide mechanistic clues as to the cellular events underpinning the progression and clinical outcomes across each phase of the IAC. Furthermore, I investigated whether an extended LS analysis, can provide an improvement of the performances of current prediction models [3, 5], that are using only three CD4+T-cells subsets.

Cohort selection and characteristic for each cohort

Participants from different clinical groups across the IAC were selected, comprising a total of 210 at-risk individuals (Pre-RA phase), 306 patients from the EAC (diagnosis phase), and 205 MXT-treated RA patients (1st line treatment phase**) (Figure 42).** A few factors led to the reduced number of patients selected for each cohort amongst the 1000's of participants enrolled in the longitudinal studies ["At-Risk CCP register" and "IACON register"]. Briefly, the B-cell panel was the most limiting amongst all participants as it was introduced late in 2015,

followed by the CD8 panel with poor marker staining limiting separation between subsets. Patients with data files for at least 4 out of the 5 panels and who also had associated clinical data were selected therefore for this project. During FC analysis of the 5 panels, more issues were encountered necessitating data to be excluded based on the 3 major exclusion rules I developed (detailed in Section **4.1.1.1, page 153**), hence limiting the at-risk cohort to 210 patients amongst the ~400 ACPA+, EAC cohort to 306 amongst the ~350, and the MXT-treated cohort to 205 amongst the ~250 participants I was able to identify from the register.

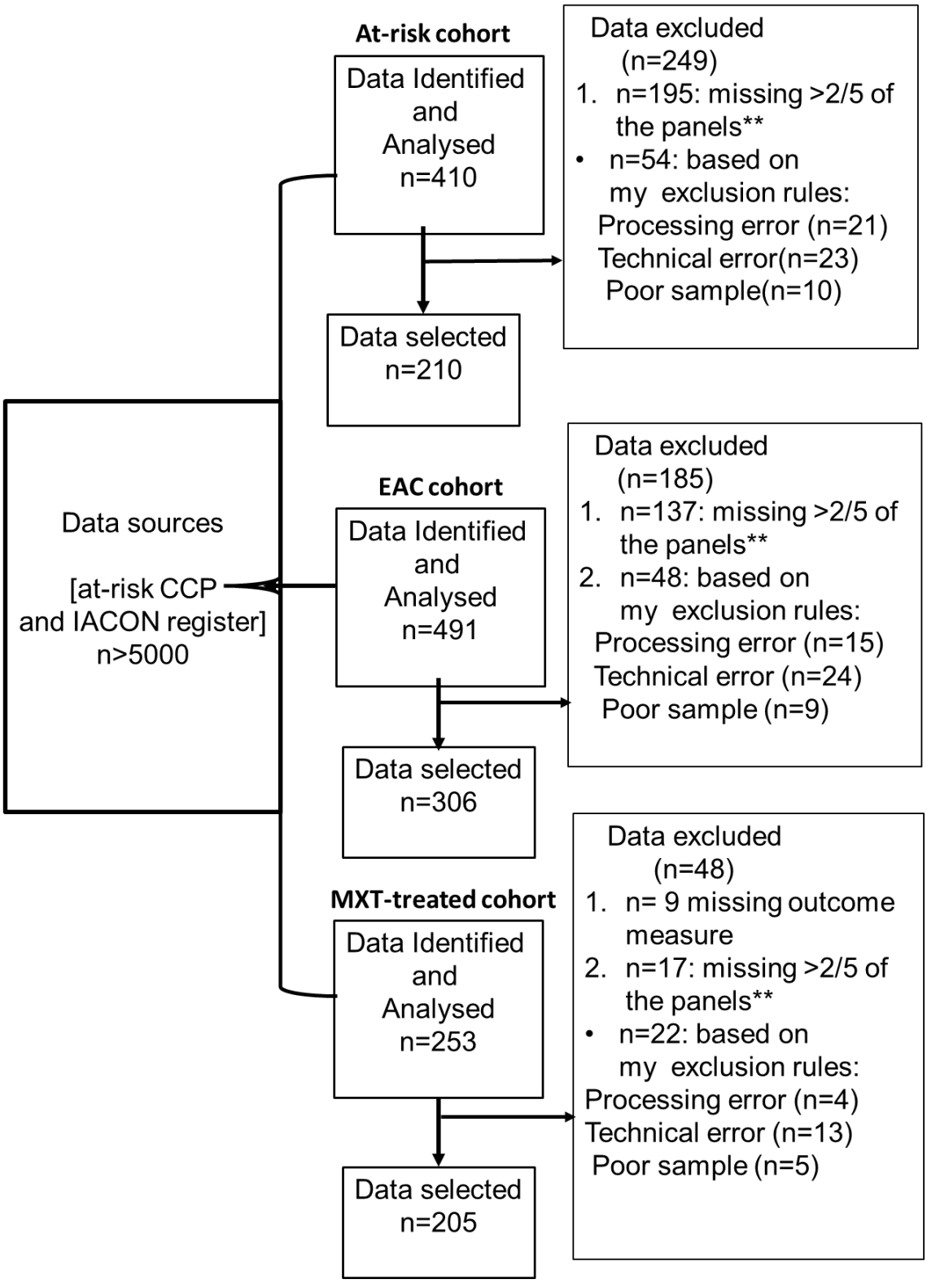


Figure 42 Patients' selection workflow across the 3 stages of the IAC.
The shows the full strategy used for selecting participants and side boxes showing reasons for exclusion at each point either based on missing LS data >2/5 panels or on exclusion criteria I developed notably due to technical issues (poor quality of blood due to transport delays). ** B- and LC- panel major limiting factors.

4.2.1.1.1 At-risk cohort outcome

Altogether, ACPA+ patients (n=210) with a minimum of 12 months of follow-up (up to 10 years), were selected from the overall cohort.

Clinical data were retrieved and progression to clinical synovitis was observed in 93/210 (44%) of participants, occurring under 12 months in 41 patients (rapid progression), within 13-24 months for 18 patients, with the last 33 progressing later than 2 years post inclusion and 1 patient after 10 years **(Figure 43**). Seventy-five percent (75%) of progressors met the EULAR 2010 Classification criteria for RA at the time of progression with an average of >3.5 swollen joints (range 1-15) and were directed to our early arthritis clinic for further care. This suggests the possibility of follow-up patients between the 3 cohorts used in the study. Unfortunately, I have no ethical permission to re-identify patients as they move from pre-clinical RA to the EAC cohort and then get treated with MTX and eventually achieve remission.

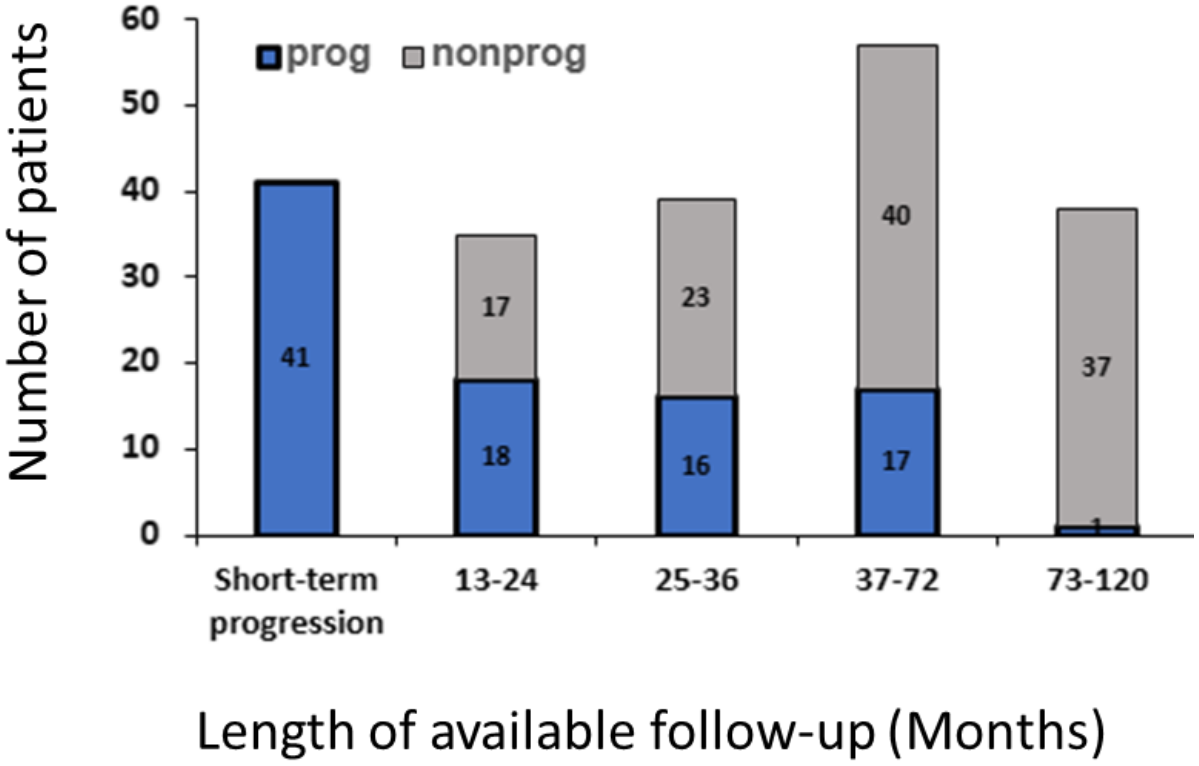


Figure 43 Length follow-up duration (Months) in progressors and non-progressors.

The length of follow-up was very wide, 120 months in total with a median follow-up duration was 34 months +/- 31 months SD overall. The shortest time to progression was 1 month while the longest was 120 months. The shortest follow-up considered in non-progressors was 12 months. The cohort was spitted between imminent progressors and then 4-time groups of delayed progression by 1 year, 2 year, 3-5 years and more than 5 years, as displayed in histogram, with BLUE shaded part showing the number of progressors and grey parted that of non-progressors with the same duration of follow-up.

Demographic/clinical data univariate association with progression

Demographic/clinical data are described in **Table 20**. Association between progression and variables at inclusion suggested 2 highly significant parameters (HLA-SE and RF, MWU $p<0.0001$ after correction) and another 3 potential (smoking, EMS,TJC78, MWU $p<0.05$), consistent with published reports in this cohort [5, 362, 520, 581]. AUCs were calculated suggesting high predictive value for RF and HLA-SE (AUC>0.650, $p<0.0001$) as well as for smoking, TJC78, EMS (AUC>0.600, $p<0.010$). The individual contribution to the prediction was however relatively small for all parameters with 19% for RF (Wald test) and >8% for the other 3 variables. Of note, no difference in ACPA-levels were detected between progressor (mean 335 OD) or non-progressors (341 OD) as well as between rapid (355 OD) and delayed (344 OD) progressors.

Table 20 Association of demographic and clinical data with progression (n=210)

	Progressors n=93 (44.3%)	Non Progressors n=117 (55.7%)	adjusted p-value[&]	AUROC (95%CI) p-value	Unadjusted OR (95%CI) p-value	Wald test
Age (years)*	53.0 (43,63)	51.0 (42,61.)	0.292	0.542 (0.464-0.620) 0.292	1.013 (0.993-1.033) 0.217	1.5
Gender (Female)	60 (64.5%)	87 (74.4%)	0.132	0.549 (0.470- 0.628) 0.221	1.595 (0.881- 0.888) 0.123	2.4
Alcohol (unit)	4.40 (0.0,10.15)	4.50 (0.0,9.80)	0.327	0.461 (0.383-0.540) 0.335	0.994 (0.974-1.014) 0.564	0.3
Smoking Never/Ever	23 (24.7%)/ 70 (75.3%)	53 (45.3%)/ 64 (54.7%)	0.002**	0.603 (0.526-0.679) 0.011	2.520 (1.390-4.571) 0.002	9.3
HLA-SE [positive]	68 (73.1%)	56 (47.9%)	<0.0001**	0.626 (0.551-0.702) 0.002	2.963 (1.651- 5.316) <0.0001	13.3
RF [Positive]	57 (61.3%)	35 (29.9%)	<0.0001**	0.657 (0.582-0.732) <0.0001	3.710 (2.087-6.593) <0.0001	19.9
Suspicion of palindromic [yes]	20 (%)	18 (%)	0.282	0.513 (0.452-0.610) 0.446	0.664 (0.328-1.343) 0.255	1.3
Family history [yes]	25 (31.6%)	32 (32.0%)	0.960	0.498 (0.413-0.584) 0.968	0.984 (0.522-1.854) 0.960	0.0
ESR (mm/hr)*	14 (6.5,20.00)	12.0 (7.0,20.5)	0.343	0.538 (0.459-0.617) 0.344	1.019 (0.993-1.045) 0.162	1.9
CRP (mg/L)*	3.180 (0.99,6.795)	3.00 (0.57,5.76)	0.159	0.557 (0.479-0.635) 0.159	1.044 (0.995-1.096) 0.082	3.0
EMS (min)*	22 (0.,60.)	5 (0,30.)	0.007[#]	0.604 (0.527-0.681) 0.009	1.007 (1.000-1.014) 0.044	4.0
TJC78*	1 (0,3)	1 (0,2.)	0.009[#]	0.601 (0.524-0.678) 0.012	1.113 (0.992-1.249) 0.068	3.3

Categorical data are presented as n (% of participant). * Numerical data are presented as median (Interquartile range values); MWU, Mann Witney U, and Chi-square tests for continues and categorical variables, respectively were used. [&]tests adjusted for 12 comparisons (adjustment of the p-value was performed by applying the Bonferroni correction method for multiple comparison tests, 0.05/12 variables = 0.0042 value or less considered significant after correction). double star**, significant; #, trend; CI, confidence interval; ESR, erythrocytes sedimentation rate; CRP, C-reactive protein; EMS, early morning stiffness HLA (SE), human leucocyte antigen (shared epitope); TJC, tender joint count; RF, rheumatoid factor; AUROC area under the roc curve.

Multivariate Modelling for progression using the clinical variable alone (Model 1)

I then used logistic regressions with a forward approach **(Table 21**) to determine the predictive value of the demographic/clinical data alone (model 1), selecting the best predictors sequentially (**Figure 44**) . Model-1 selected 4 parameters in a stepwise construction, starting with the RF as shown in the step below

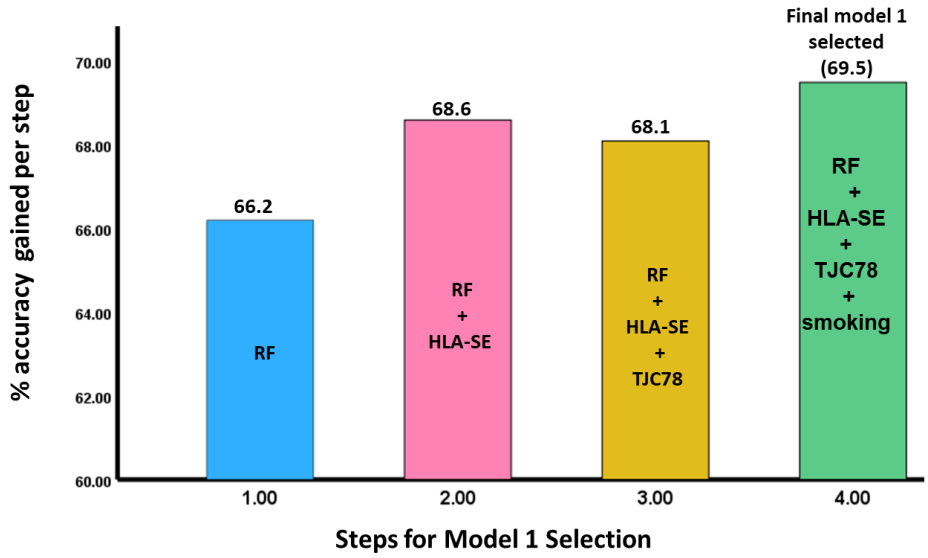


Figure 44. Accuracy gains in logistic regression modelling of Clinical data (Model 1):
The histogram illustrated the steps and order of variables selection. The blue bar represents step 1, the purple bar represents step 2, Yellow bar represents step 3 and green bar shows the final step retaining 4 variables. Model accuracy gain at each step is indicated on each bar.

This model accurately predicted 70% of cases, with Sensitivity and specificity=56% and 81.2% and a good NPV and PPV=70%, with an AUC=0.744 **(Table 21).** However, only 25% of the variance for predicting progression was accounted for (Nagelkerke R-square) and individual variables contributed 17% for RF and less than 7% for the other 3 (individual Wald score).

Table 21 Modelling for the predicting of overall progression with clinical variables only

	Logistic regression OR (95% CI) p-value (Wald test) n=210	
	unadjusted	**Model-1**
Smokers (ever)	2.520 (1.390-4.571) 0.002 (9.3)	2.282 (1.1.187-4.388) 0.013 (6.1)
HLA-SE positive	2.963 (1.651- 5.316) <0.0001 (13.3)	2.527 (1.335-4.782) 0.004 (8.1)
RF positive	3.710 (2.087-6.593) <0.0001 (19.9)	3.600 (1.947-6.656) <0.0001 (16.7)
TJC78	1.113 (0.992-1.249) 0.068 (3.3)	1.178 (11.039-1.336) 0.010 (6.6)
Accuracy (%)	not applicable	70.0%
AUROC (95%CI) p-value		0.744 (0.678-0.810) p<0.0001
Sensitivity (%) (95%CI)		55.9 (55-66)
Specificity (%) (95%CI)		81.2 (73-87)
PPV (%) (95%CI)		70 (61-78)
NPV (%) (95%CI)		70 (64-75)
Nagelkerke R square		25%
Hosmer& Lemeshow		0.853

ESR, erythrocytes sedimentation rate; CRP, C-reactive protein; EMS, early morning stiffness HLA (SE), human leucocyte antigen (shared epitope); TJC, tender joint count; RF, rheumatoid factor; OR, odd ratio; AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. CI, Confidence interval. NB: Table showed all variables retained in the model.

Flow cytometry analysis and univariate association with progression

Fresh Blood samples at inclusion in the study were sent from clinics at Chapel-Allerton Hospital, to the NHS immunology services at the St James hospital using NHS transport services. Blood LS data were acquired by NHS staff within a few hours of the samples arriving at the NHS immunology lab, with some exceptional delays on some occasion. The raw data (FCS files) were made available for use in my PhD project as electronic files.

Data missingness and imputation

FC data analysis was performed and data recorded. All selected patients (n=210) had data for the CD4+T-cells panel (naïve, IRC) and the lymphocyte count panel. For the Treg panel, shortage of 1 antibody (FoxP3) led to missing data over a few weeks (n=20/210). For the CD8 panel, I had issue with poor separation between markers and difficulties in gating the CD8-IRC subset. This was associated with samples for which processing was delayed by transport. Furthermore, faulty acquisition due to non-adherence to sample preparation protocols or problems encountered over data acquisition (for example the flow machine breaking down), led to more data being excluded .

Overall, 150/210 patients had a full dataset (i.e.. for all 18 subsets analysed) and I imputed data for individual subsets in 41 patients where I had at least >15/18 LS present. I further imputed data for the B-panel in 19 patients who had complete data for all other LS. Missing data imputation was performed using the SPSS "Random Number Generators" function and then the "Imputing Missing Data Values" function over 5 cycles. The estimates obtained from each dataset (5 imputation datasets) were aggregated to produce an overall multiple imputation estimate using the same SPSS package. The pooled dataset after imputation

was compared to the non-imputed dataset to verify that OR were not affected by the imputation process using SPSS. Diagnostic checks [731-734] confirmed the validity of the imputed dataset, showing no significant differences compared to the pre-imputed dataset **(Table 22).** This indicates that the imputed data accurately reflected the original dataset.

Table 22 Data distribution characteristics of original dataset compared with imputed dataset for the at-risk cohort

Cell subsets	Original data Median (IQR)	Imputed data Median (IQR)	MWU p-value
CD4 T-cells	51.5 (45.88, 56.98)	51.49 (46.08, 56.34)	0.919
CD8 T-cells	19.01 (15.74, 24.45)	19.72 (15.76, 24.51)	0.897
B-cells	11.59 (9.58, 14.48)	11.67 (9.41, 14.54)	0.848
NK cell sCD56^{bright}	0.46 (0.33 to 0.62)	0.42 (0.29, 0.56)	0.865
NK cells CD56^{dim}	8.42 (6.58, 11.75)	8.68 (6.83, 11.76)	0.837
NKT	1.52 (0.93, 3.29)	1.82 (0.95, 3.64)	0.766
naive CD4 cells$^{\$}$	4.79 (-5.91 ,11.05)	4.12 (-5.02, 11.09)	0.973
IRCCD4 cells	1.02 (0.30, 2.12)	1.40 (0.50 to 2.65)	0.884
memory CD4 cells$^{\$}$	-8.1 (-11.92, -2.07)	-8.1 (-11.29, -2.85)	0.953
Treg CD4 cells $^{\$}$	-0.27 (-1.2 , 1.26)	-0.38 (-1.47, 0.94)	0.812
NaïveCD8 cells$^{\$}$	7.14 (-1.33, 14.78)	7.39 (0.58, 14.26)	0.910
MemoryCD8 cells$^{\$}$	-23.4 (-30.02, -4.49)	-23.52 (-29.22, -15.53)	0.989
exp-mem CD8 cells	10.2 (5.81, 16.10)	9.72 (6.25, 14.09)	0.791
IRC CD8 cells	10.3 (6.40, 19.90)	11.97 (6.80, 20.24)	0.850
Naive B-cells$^{\$}$	-0.91 (-10.46, 7.77)	-0.91 (-10.49, 6.89)	0.937
Memo B-cells $^{\$}$	1.59 (-8.47, 11.14)	1.47 (-6.74 ,10.69)	0.977
B-reg	0.54 (-0.65, 2.04)	0.8 (-0.20, 2.14)	0.807
PBs	0.5 (0.2, 0.9)	0.5 (0.3, 0.9)	0.858

Data are presented as median (1st and 3rd Interquartile values); CI Confidence interval; $ normalised subsets, NK natural killer, NKT natural killer-T, Treg regulatory T-cells, IRC, inflammatory-related cells; & MWU, Mann Witney U test . Breg= Regulatory B-cells, PBs= Plasmablasts

The overall results of LS quantification **(Figure 45, Table 23)** showed a significant association with progression for 5 subsets:

- lower naïve CD4+T-cells ($p<0.0001$)
- Lower Treg ($p<0.0001$)
- lower CD8+T-cells ($p=0.021$)
- higher CD4-IRC ($p<0.0001$)
- Higher B-reg ($p=0.015$)

Three more subsets showed non-significant difference after correction, that may nonetheless suggest other biological events leading to progression.

- higher NK-cells $CD56^{bright}$ ($p=0.026$ before/$p=0.154$ after)
- higher memory B-cells ($p=0.023$ before/$p=0.135$ after)
- lower plasmablasts ($p=0.018$ before/$p=0.105$ after)

AUC calculated and these 8 LS showed significant predictive values/trend for progression. Of note, Treg, naïve CD4+T-cells and IRC showed the highest contribution to the prediction (Table 23, Wald score 27%, 15% and 12% respectively) while >6% each for CD8, CD56, Breg and plasmablasts. These subsets had high specificity for progression (all >80%) but relatively low sensitivities (32-40%), except for CD4+Treg (60%) (**Table 24)**.

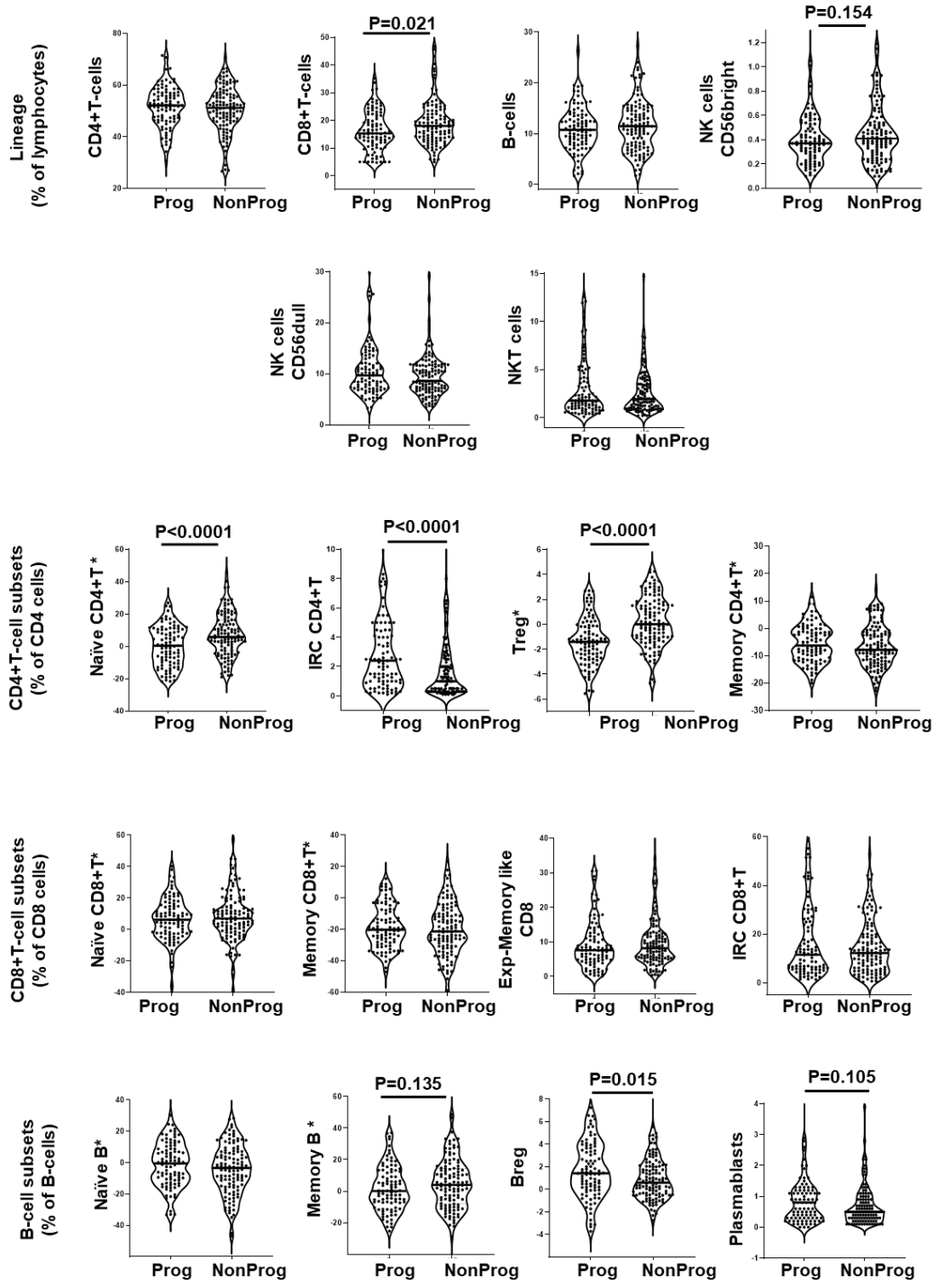


Figure 45 Frequency of lineage and lymphocytes subsets in at-risk progressors vs non-progressors.

LS were analysed by flow cytometry and data displayed as violin plots (each dot representing a patient) for Progressor (Prog, n=93) and non-Progressors (Non-Prog, n=117). Star (*) indicate LS that were normalised as previously described [5]. P-value corrected for multiple testing (MWU test) are indicated when significant and # designate trends.

Table 23 Association of 18 LS with Progression (n=210)

	Missing Data n (%)	Progressors n=93 (44.3%)	Non Progressors n=117 (55.7%)	adjusted p-value[&]	AUROC (95%CI) p-value	Unadjusted OR (95%CI) P value	Wald test (%)
CD4 T-cells	none	52.11 (46.41 56.96)	51.18 (45.98 56.35)	0.479	0.528 (0.450-0.607) 0.479	1.014 (0.982-1.048) 0.394	0.7
CD8 T-cells	19 (9%)	15.34 (11.57 21.97)	17.98 (13.68 23.46)	0.021	0.407 (0.330-0.485) 0.021	0.952 (0.917-.988) 0.009	6.8
B cells	none	10.73 (7.97 13.55)	11.44 (7.41 14.66)	0.503	0.473 (0.395-0.551) 0.503	0.972 (0.917-1.030) 0.337	0.9
NK cells CD56^{bright}	none	0.37 (0.27 0.53)	0.41 (0.27 0.57)	0.154[#]	0.447 (0.369-0.524) 0.164	0.485 (0.178-1.317) 0.156	2.0
NK cells CD56^{dim}	none	9.69 (7.10 12.53)	8.61 (6.85 11.67)	0.207	0.551 (0.472-0.629) 0.205	1.052 (0.987-1.120) 0.117	2.5
NKT cells	none	1.78 (1.01 4.16)	1.93 (0.96 3.87)	0.878	0.506 (0.427-0.586) 0.878	1.021 (0.943-1.106) 0.610	0.2
Naïve CD4 cells [$]	13 (6.5%)	0.42 (-10.10 10.58)	5.74 (-2.59 16.38)	**<0.0001***	0.355 (0.280-0.429) <0.0001	0.956 (0.934-0.978) <0.0001	14.8
Memory CD4 cells[$]	23 (11%)	-6.23 (-11.34 -0.38)	-7.81 (-11.93 - 1.17)	0.327	0.539 (0.461-0.618) 0.327	1.016 (0.979-1.055) 0.407	0.7
IRC CD4 cells	13 (6.5%)	2.40 (1.00 4.50)	1.00 (0.30 2.50)	**<0.0001***	0.658 (0.584-0.732) <0.0001	1.245 (1.098- 1.411) <0.0001	11.7
Treg CD4 cells[$]	20 (9.3)	-1.41 (-2.71 -0.13)	0.01 (-1.10 1.64)	**<0.0001***	0.273 (0.205-0.342) <0.0001	0.651 (0.554-0.765) <0.0001	27.0
Naïve CD8 cells [$]	35 (16.5%)	6.16 (-2.60 14.78)	6.91 (-0.82 14.78)	0.448	0.469 (0.391-0.548) 0.448	0.992 (0.973-1.011) 0.387	0.7
Memory CD8 cells [$]	36 (17%)	-20.22 (-28.92 -7.65)	-21.32 (-30.83-9.72)	0.533	0.525 (0.477-0.603) 0.533	1.006 (0.988-1.024) 0.516	0.4
Exp-memory like CD8 cells	39 (19%)	7.60 (4.46 12.40)	8.20 (5.32 12.40)	0.296	0.458 (0.379-0.537) 0.296	0.980 (0.945-1.017) 0.281	1.2
IRC CD8 cells	37 (18%)	11.50 (5.55 25.09)	12.20 (5.60 19.90)	0.959	0.498 (0.418-0.578) 0.959	1.006 (0.986-1.027) 0.536	0.4
Naïve B cells [$]	19 (12.6%)	-0.50 (-11.19 8.83)	-3.39 (-15.10 6.94)	0.200	0.554 (0.476-0.632) 0.189	1.014 (0.995-1.033) 0.179	2.2
Memory B cells[$]		0.10 (-7.35 12.42)	4.05 (-7.89 14.42)	0.135[#]	0.440 (0.362-0.518) 0.138	.986 (0.967-1.004) 0.125	2.4
Breg		1.42 (-0.05 3.56)	0.60 (-0.47 2.05)	**0.015***	0.598 (0.520-0.677) 0.015	1.164 (1.021-1.325) 0.021	5.3
PBs		0.80 (0.30 1.20)	0.50 (0.30 0.95)	0.105[#]	0.565 (0.486-0.644) 0.106	1.119 (0.873-1.432) 0.175	0.8

Data are presented as median (1st and 3rd Interquartile values); CI Confidence interval; $ normalised subsets, NK natural killer, NKT natural killer-T, Treg regulatory T-cells, IRC, inflammatory-related cells; Exp-, expanded, & MWU Mann Witney U test adjusted p-value for 18 comparison (adjustment of the p-value was performed by applying Bonferroni correction method for multiple comparison test, AUROC, area under the roc curve. P-value bold with (*) indicated significant value and with (#) indicated trend. Breg= Regulatory B cells, PBs= Plasmablasts

Table 24 Univariate performance index of the 18 subsets in relation to association with predicting progression to IA

	SENSTIVITY (95%CI)	SPECIFICTY (95%CI)	PPV (95%CI)	NPV (95%CI)	overall % accuracy (95%CI)
CD4 T-cells	20.51 (13.61 to 28.97)	80.12 (73.34 to 85.82)	41.38 (30.68 to 52.95)	59.57 (56.68 to 62.38)	55.90 (49.96 to 61.72)
CD8 T-cells	32.48 (24.11 to 41.76)	80.12 (73.34 to 85.82)	52.78 (42.87 to 62.47)	63.43 (59.97 to 66.75)	60.76 (54.86 to 66.44)
B cells	21.37 (14.33 to 29.91)	80.70 (73.98 to 86.33)	43.10 (32.28 to 54.63)	60.00 (57.10 to 62.83)	56.60 (50.66 to 62.40)
NK cells CD56^{bright}	23.93 (16.53 to 32.70)	80.12 (73.34 to 85.82)	45.16 (34.62 to 56.15)	60.62 (57.57 to 63.59)	57.29 (51.35 to 63.08)
NK cells CD56^{dim}	30.77 (22.57 to 39.97)	80.12 (73.34 to 85.82)	51.43 (41.38 to 61.36)	62.84 (59.47 to 66.10)	60.07 (54.16 to 65.77)
NKT cells	26.50 (18.77 to 35.45)	80.12 (73.34 to 85.82)	47.69 (37.32 to 58.27)	61.43 (58.27 to 64.51)	58.33 (52.40 to 64.09)
Naïve CD4 cells	39.32 (30.41 to 48.77)	80.12 (73.34 to 85.82)	57.50 (48.16 to 66.33)	65.87 (62.09 to 69.45)	63.54 (57.69 to 69.11)
Memory CD4 cells$	25.64 (18.02 to 34.54)	80.12 (73.34 to 85.82)	46.88 (36.44 to 57.59)	61.16 (58.03 to 64.20)	57.99 (52.05 to 63.75)
IRC CD4 cells	39.32 (30.41 to 48.77)	80.12 (73.34 to 85.82)	57.50 (48.16 to 66.33)	65.87 (62.09 to 69.45)	63.54 (57.69 to 69.11)
Treg CD4 cells$	59.83 (50.36 to 68.78)	80.12 (73.34 to 85.82)	67.31 (59.55 to 74.22)	74.46 (69.77 to 78.64)	71.88 (66.30 to 76.99)
Naïve CD8 cells $	23.93 (16.53 to 32.70)	80.12 (73.34 to 85.82)	45.16 (34.62 to 56.15)	60.62 (57.57 to 63.59)	57.29 (51.35 to 63.08)
Memory CD8 cells $	32.48 (24.11 to 41.76)	80.12 (73.34 to 85.82)	52.78 (42.87 to 62.47)	63.43 (59.97 to 66.75)	60.76 (54.86 to 66.44)
Exp-memory like CD8 cells	30.77 (22.57 to 39.97)	78.95(72.07 to 84.80)	50.00 (40.19 to 59.81)	78.95 (72.07 to 84.80)	59.38 (53.46 to 65.10)
IRC CD8 cells	23.08 (15.79 to 31.77)	80.12 (73.34 to 85.82)	44.26 (33.68 to 55.40)	60.35 (57.35 to 63.28)	56.94 (51.01 to 62.74)
Naïve B cells $	22.22 (15.06 to 30.84)	80.12 (73.34 to 85.82)	43.33 (32.71 to 54.61)	60.09 (57.12 to 62.98)	56.60 (50.66 to 62.40)
Memory B cells$	23.93 (16.53 to 32.70	79.53 (72.70 to 85.31)	44.44 (34.05 to 55.35)	60.44 (57.37 to 63.44)	56.94 (51.01 to 62.74)
Regulatory B cells	35.90 (27.24 to 45.29)	80.12 (73.34 to 85.82)	55.26 (45.64 to 64.51)	64.62 (61.01 to 68.08)	62.15 (56.28 to 67.78)
Plasmablasts	38.46 (29.62 to 47.91)	81.29 (74.62 to 86.83)	58.44 (48.84 to 67.44)	65.88 (62.19 to 69.38)	63.89 (58.05 to 69.44)

Data are presented as median (1st and 3rd Interquartile values); CI Confidence interval; $ normalised subsets, NK natural killer, NKT natural killer-T, Treg regulatory T-cells, IRC, inflammatory-related cells; . Breg= Regulatory B cells, PBs= Plasmablasts PPV, positive predictive value; NPV, negative predictive value. CI, Confidence interval**.**

Clusters of patients

To decipher whether group of patients had similar immune cell profiles, an unsupervised hierarchical clustering algorithm was applied to log transformed frequencies for 18 LS (n=210) to define groups of patients with similar dynamic change in LS profiles while not specifying the clinical outcome. This clustering algorithm builds relationships between LS frequencies based on spearman rank correlations. First analysis using the 18 LS displayed as a heat-map of frequencies **(Figure 46),** showed 3 distinct profiles of the LS. The 1st group included 5 subsets which frequency reduced with age in health (naïve CD4/CD8 T-cells and memory B-cells) as well as CD4+T-cells and the expanded-CD8 subset. The 2nd group clustered 6 subsets including those increasing with age (memory CD4/CD8 T-cells, Treg) as well as CD8+T-cells, CD8-IRC and PBs. The 3rd group gathered 7 subsets: B-cells, naïve B-cells and Breg, as well as NK (CD56 bright and dim), NKT-cells and CD4-IRC. Therefore, each group contained at least 1 of the highly predictive subsets of progression [1st group (naïve CD4), 2nd group (Treg), 3rd group (Breg)], suggesting that these were dysregulated independently of each other.

The analysis also segregated patients into 3 clusters (annotated I, 2 and 3), based on 3 main group of LS. One with 90 patients, a second with 38 patients and a third with 82 patients . Cluster of patients showed significantly different proportions of progressors (Chi-square test, $p<0.0001$) with 39/90 (45%) progressors in the cluster-1, depleted with 5/38 (13%) in cluster-2, and enriched in cluster-3 with 49/82 (60%), suggesting that different LS profiles may be able to discriminate patients with different outcomes.

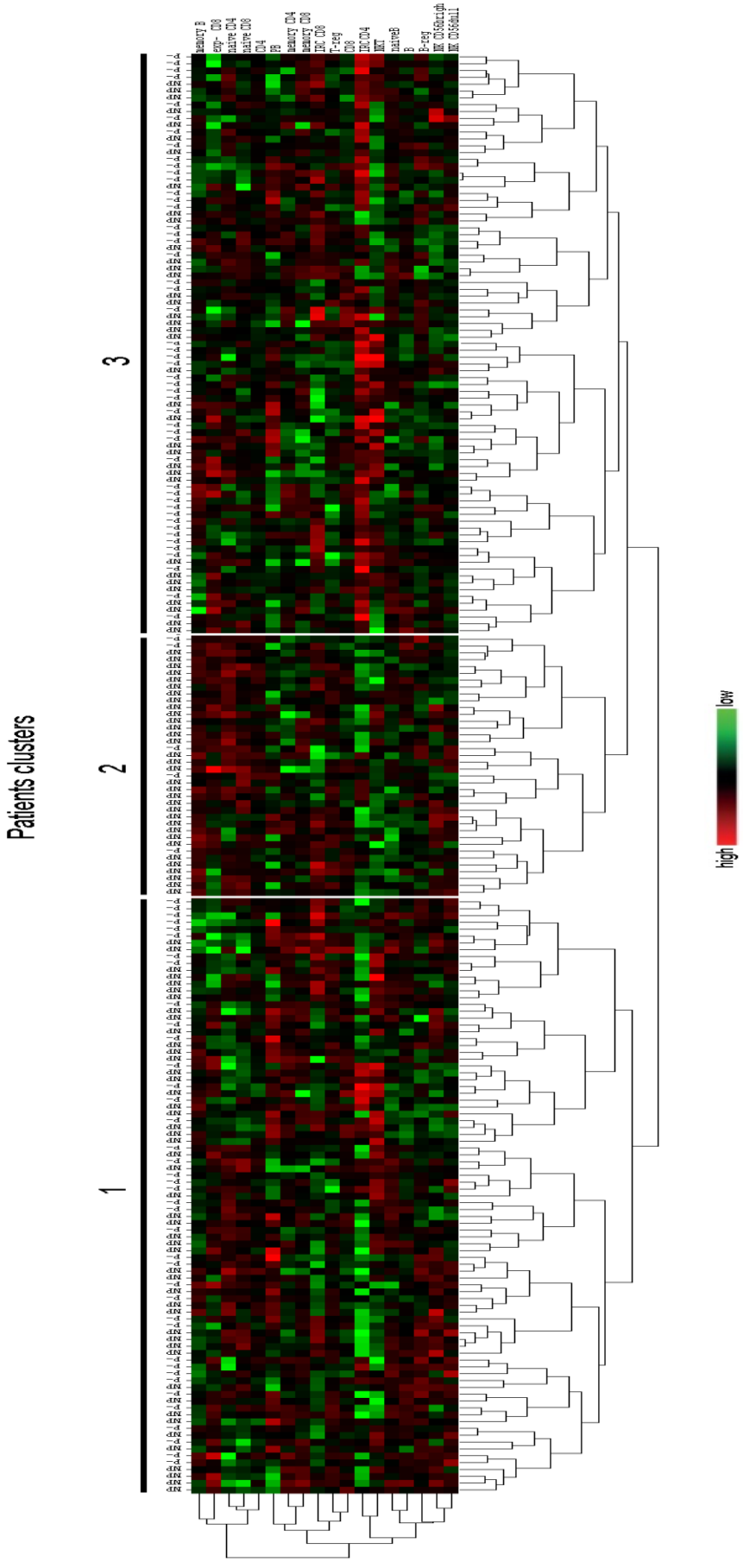

Patients clusters
1
2
3
memory B
exp- CD8
naive CD4
naive CD8
CD4
NK
memory CD4
memory CD8
IRC CD8
T-reg
CD8
IRC CD4
MAIT
naiveB
B
P-reg
NK CD56bright
NK CD56dull
low
high

Figure 46 Unsupervised hieratical clustering of 18 subsets in individual at-risk of developing IA (n=210).

Heat-map display showing three cluster of patient annotated with 1, 2 and 3 with the highest LS frequency (red) and (green) the lowest frequency observed in each cluster. Dendrogram (bottom) represented 3 LS profile group. NP= Non-progressors, P= progressors. Chi-square test, $p<0.0001$ for progressors between patient clusters

I repeated this approach limiting the data used to the 8 subsets identified as more likely to be predictive. This analysis distributed LS in 4 groups of LS in 3 clusters of patients (

Figure 47**).** This is further illustrated in **figure 48**, highlighting the subset imbalance.

- The 1st subset LS included only plasmablasts,
- the 2nd included naïve CD4+T, B-reg and NK-CD56^{bright} cells;
- the 3rd combined memory B-cells, Treg and CD8+T-cells.
- The last group with IRC-CD4 alone.

For patients, Cluster-1 (n=60) was driven by high Treg and high naïve CD4 but very low IRC-CD4 and mixed NKCD56bright and was mainly composed of non-progressors 47/60 (78%).

Cluster-2 (n=41) was mainly driven by high plasmablasts while Treg were also high and CD4+IRC low. This profile was associated with a mix outcome of 17/41 (42%) progressors and 24/41 (58%) non-progressors.

In cluster-3, the largest (n=109), IRC-CD4 was particularly high and all other LS showed mixed patterns defining subgroups of patients. The proportion of progressors in cluster-3 was 58% (63/109), which was significantly higher than in the other 2 clusters (Chi-square test, $p<0.0001$).

There was no difference in any demographic or clinical data between these 3 clusters of patients. I, however, observed higher levels of ACPA measured by a 2nd generation CCP-2 test (**Figure 49**, p=0.050, mean 220 OD in cluster 3 compared to cluster 1 with 158 OD, but not with cluster 2 (253 OD, p=0.177). No association was seen between plasmablasts and ACPA (or RF) levels, notably in Cluster 2 showing the highest plasmablasts frequencies.

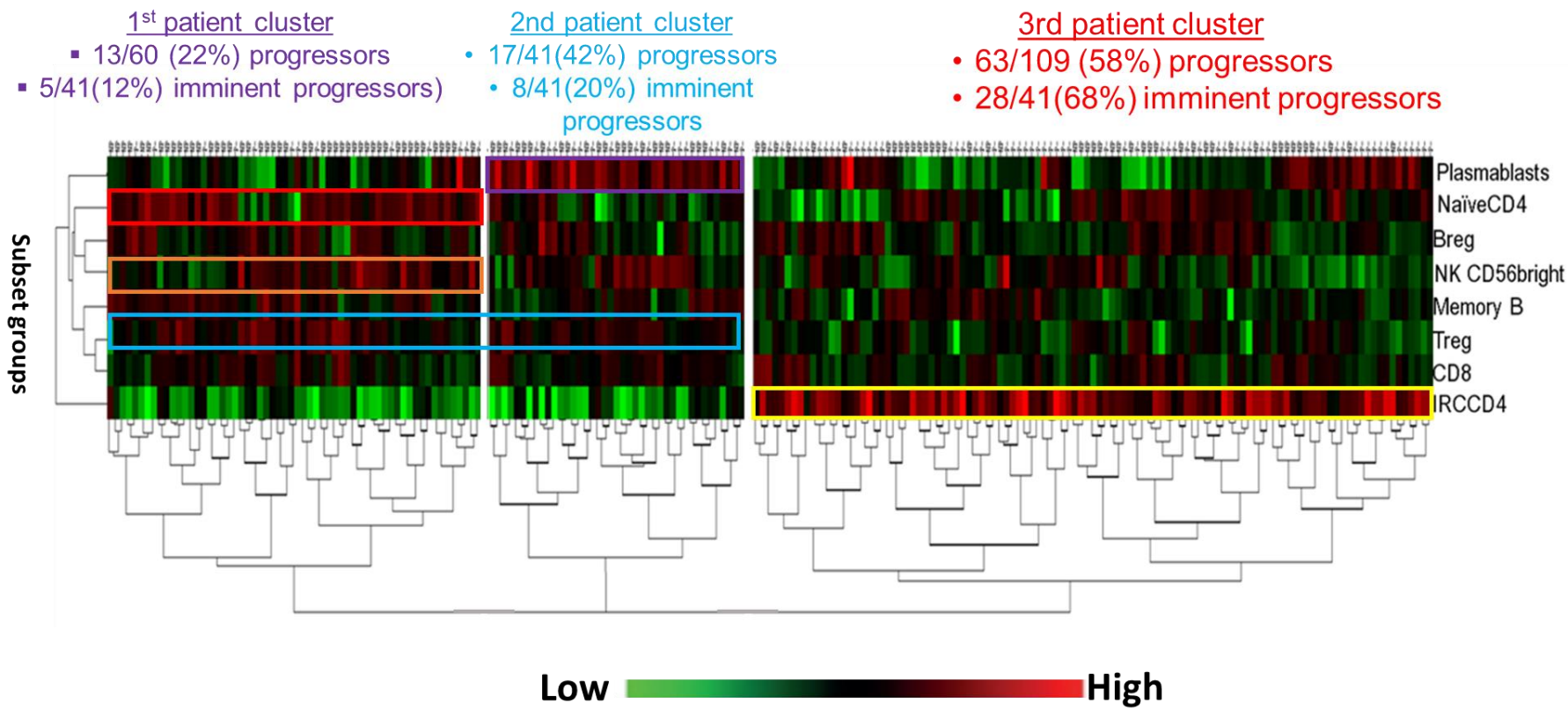

1st patient cluster
13/60 (22%) progressors
5/41(12%) imminent progressors)
2nd patient cluster
17/41(42%) progressors
8/41(20%) imminent progressors
3rd patient cluster
63/109 (58%) progressors
28/41(68%) imminent progressors
Subset groups
Plasmablasts
NaïveCD4
Breg
NK CD56bright
Memory B
Treg
CD8
IRCCD4
Low
High

Figure 47 Unsupervised hierarchical clustering of the 8 subsets associated with progression to IA (n=210).

An unsupervised hierarchical clustering algorithm was applied to log transformed frequencies for 8 LS and results are displayed as a heat-map of data (red being the highest and green the lowest frequency observed for each LS). This clustering algorithm builds relationships between LS frequencies based on spearman rank correlations, and segregated patients into 3 clusters (1, 2 and 3), annotated with the % of progressors. The first group of LS (plasmablasts) shows particular high frequencies in patient cluster-2 (purple) (mix of progressors and non-progressors). The 2nd group with 2 subsets [(naiveCD4(red) and NK CD56bright (orange)] defined Cluster-1 (mostly non-progressors). The third group (with Treg) allows to define both Cluster 1 and 2(blue). The last group (IRC-CD4 (yellow)) shows exclusively high frequencies in cluster-3 (enriched progressors) with lower frequencies in Cluster-1 and 2. The proportion of progressors to IA in the 3 clusters was significantly different ($p<0.0001$).

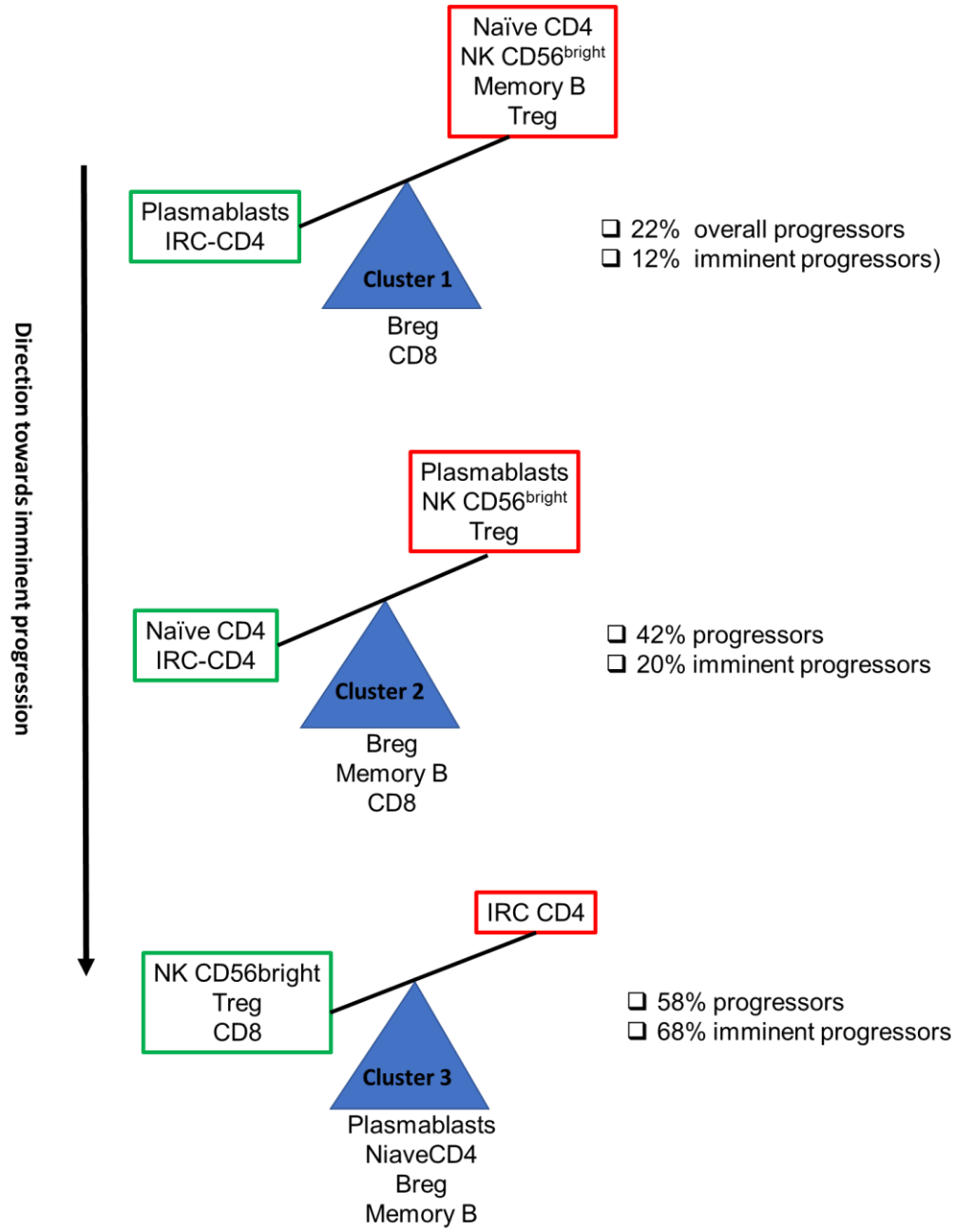


Figure 48 Scale diagram illustrate subsets imbalance driving progression from the cluster analysis.
Subsets with lower frequencies(green box), higher frequencies(red box), and those with mix pattern of frequencies(under the base of the scale)

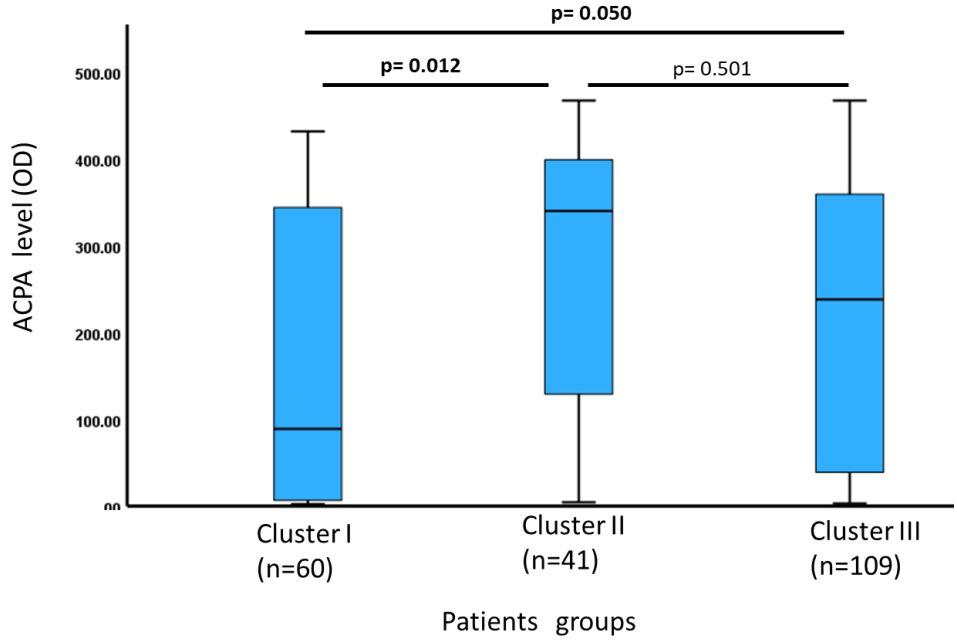


Figure 49 ACPA level across the patient's cluster.
Distribution of ACPA level in each cluster presented in box plots, Median value (dark line within the box), Median value = 89 OD (Cluster I), 340 OD (cluster II), and 238 OD (cluster III). Significant p-value indicated in bold

Multivariate modelling for the prediction of progression across the IAC

A predictive value for 3 CD4+T-cell subsets (naïve, IRC, Treg) was previously reported [5]. Modelling for progression to IA using an enter approach in these 210 patients (as in previous work, **Table 25, Model 3**) [5], confirmed previous data with an accuracy=78.6% with an AUC=0.880, although IRC did not independently contribute to this model (p=0.144).

This model 3 showed definite better performance than a model using the same data/strategy but using only clinical data (Model 1, 69.6%, AUC= 0.744) or the 3 LS (Model 2, 71.9%, AUC=0.880). This was already demonstrating the added value of these 3 CD4+T subsets.

To demonstrate both the predictive value and the added value of 18 LS quantification, I then performed different modelling. Due to the amount of variables, I decided to adopt a forward logistic strategy, allowing to select the best combination of predictors amongst a large list of 18 LS.

Table 25 Three CD4+T-cell subsets modelling (n=210): unadjusted and multivariate logistic regression using an enter method as previously reported (15)

	Logistic regression OR (95% CI) p-value			
	unadjusted	**Model-1**	**Model-2**	**Model-3**
Smokers (ever)	2.520 (1.390-4.571) 0.002	2.282 (1.1.187-4.388) **0.013**		4.020 (1.740-9.290) **0.010**
HLA-SE positive	2.963 (1.651- 5.316) <0.0001	2.527 (1.335-4.782) **0.004**		2.747 (1.256-6.007) **0.011**
RF positive	3.710 (2.087-6.593) <0.0001	3.600 (1.947-6.656) **<0.0001**		3.776 (1.773-8.040) **0.0001**
TJC78	1.113 (0.992-1.249) 0.068	1.178 (11.039-1.336) **0.010**		1.216 (1.030-1.435) **0.021**
Naïve CD4 cells $	0.956 (0.934-0.978) <0.0001		0.932 (0.905-0.952) **<0.0001**	0.926 (0.898-0.957) **<0.0001**
IRC CD4 cells	1.245 (1.098- 1.411) <0.0001		1.099 (0.958-1.232) 0.177	1.104 (0.955-1.300) 0.144
Treg CD4 cells$	0.651 (0.554-0.765) <0.0001		0.578 (0.476-0.704) **<0.0001**	0.528 (0.420-0.662) **<0.0001**
Accuracy (%)		**69.5%**	**71.9%**	**78.6%**
AUROC (95%CI) p-value		0.744 (0.678-0.810) p<0.0001	0.807 (0.714-0.865) p<0.0001	**0.880** (0.831-0.921) <0.0001
Sensitivity (%) (95%CI)		55.9 (55-66)	62.3 (52-72)	**73.1** (63-82)
Specificity (%) (95%CI)		81.2 (73-87)	74 (65-81)	**82.9** (75-89)
PPV (%) (95%CI)		70 (61-78)	63 (55-70)	**77.3** (69-84)
NPV (%) (95%CI)		70 (64-75)	73.5 (68-78.5)	**79.5** (73-84)

HLA (SE), human leucocyte antigen (shared epitope); TJC, tender joint count; RF, rheumatoid factor; Treg, regulatory T-cells; IRC, inflammatory-related cells; $ normalised frequency, OR, odd ratio; AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. CI, Confidence interval.

Modelling for the LS alone (model-2)

Modelling with 18 LS with logistic regression **(Table 26)** excluded many subsets and retained only 5 at the end with steps below selecting the best predictor (Treg) followed by those that improved the model's performance **(Figure 50**). An additional step adding, plasmablasts (although not independently predictive, table 6, p=0.085) increased accuracy by ~1% with no improvement of AUC=0.862. As such the final model was selected at step 5

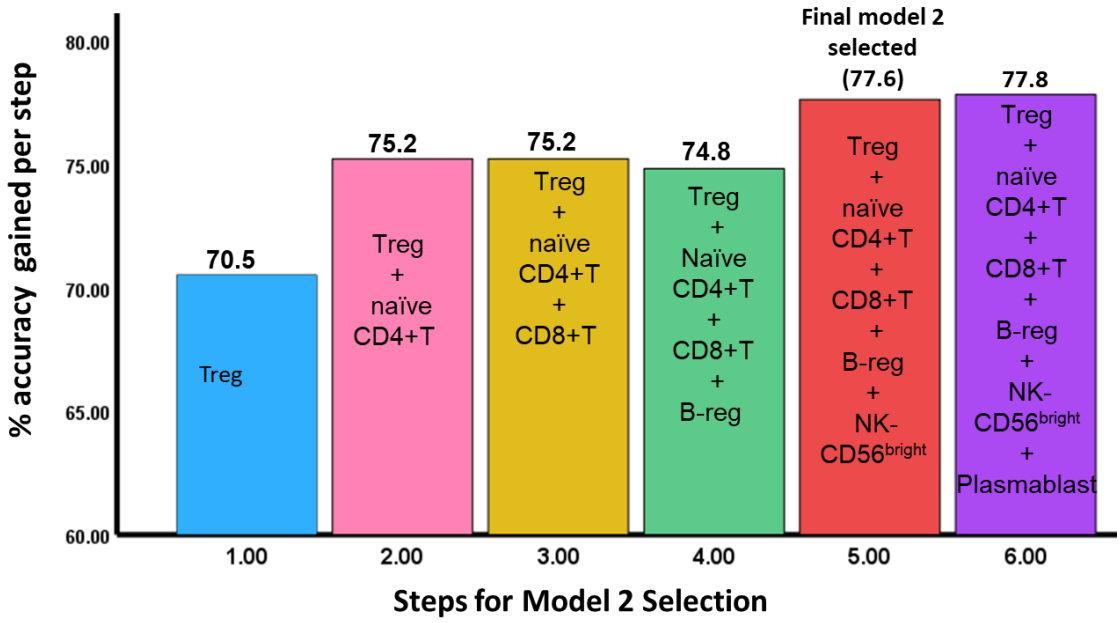


Figure 50 Accuracy gains in logistic regression modelling of LS data (Model 2):
The histogram illustrated the steps and order of variables selection. The blue bar represents step 1, the purple bar represents step 2, the yellow bar represents step 3 and so on till step 5 (red bar) the retaining 5 variables. Model accuracy gain at each step is indicated on each bar.

Model-2 accounted for 50% of the variance with Treg (34%), naïve CD4+T-cells (32%) and CD8 (14%) contributing with the highest and other LS each for >6%. Altogether model 2 had better accuracy than Model 1.

Table 26 Modelling for the predicting of overall progression with 18 LS

	Logistic regression OR (95% CI) p-value (Wald test) n=210	
	unadjusted	**Model-2**
CD8 T-cells	0.952 (0.917-.988) 0.009 (6.8)	0.911 (0.867-0.957) <0.0001 (13.7)
NK cells CD56^{bright}	0.485 (0.178-1.317) 0.156 (2.0)	0.155 (0.038-0.631) 0.009 (6.8)
Naïve CD4 cells $	0.956 (0.934-0.978) <0.0001 (14.8)	0.899 (0.867-0.932) <0.0001 (32.6)
Treg CD4 cells$	0.651 (0.554-0.765) <0.0001 (27.0)	0.518 (0.416-0.646) <0.0001 (34.2)
Regulatory B cells	1.164 (1.021-1.325) 0.021 (5.3)	1.253 (1.061-1.479) 0.008 (7.0)
Plasma blasts	1.119 (0.873-1.432) 0.375 (0.8)	1.326 (0.962-1.826) 0.085 (2.9)
Accuracy (%)	not applicable	77.6%
AUROC (95%Cl) p-value		0.862 (0.814-0.910) p<0.0001
Sensitivity (%) (95%Cl)		74.2 (64-83)
Specificity (%) (95%Cl)		80 (72-87)
PPV (%) (95%Cl)		75 (67-83)
NPV (%) (95%Cl)		80 (73-85)
Nagelkerke R square		50%
Hosmer& Lemeshow		0.912

NK, natural killer; NKT, natural killer-T; Treg, regulatory T-cells; IRC, inflammatory-related cells; Exp-, expanded, $ normalised frequency, OR, odd ratio; AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. CI, Confidence interval. NB; Table presented only LS retained in the model

Clinical and LS modelling (model 3).

To develop a model accounting for both clinical variables on LS in predicting progression, I performed modelling combining both datasets **(Table 27).** Model-3 was developed on 10 steps **(Figure 51**) selecting 4 clinical and 6 LS. Again the best LS was Treg, then naïve. Clinical data entered the model as 3rd/4th step and the final step still increased accuracy with the last entry (plasmablasts) were not fully providing independent contribution (Table 7, p=0. 0.07).

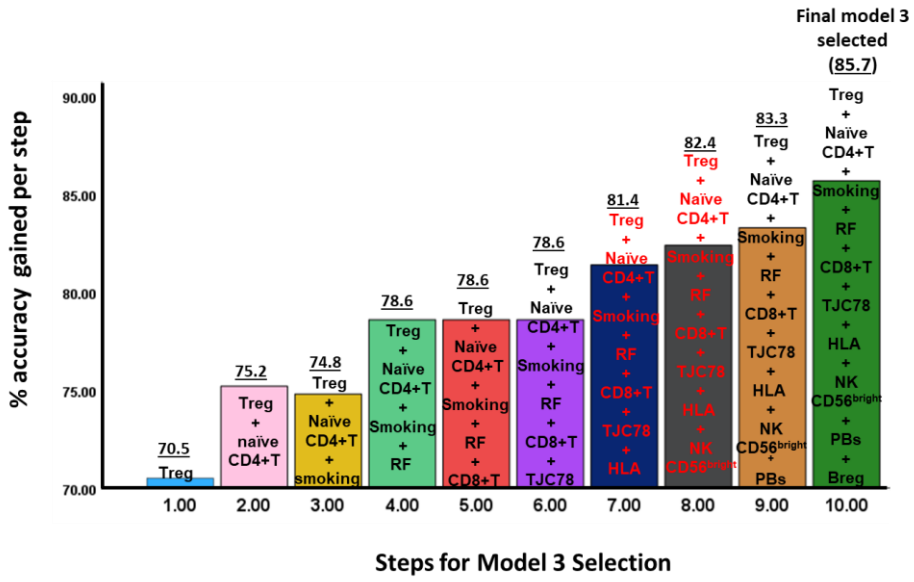


Figure 51 Accuracy gains in logistic regression modelling of Clinical and LS (Model 3):
The histogram illustrated the steps (1-10) and order of variables selection. Model accuracy gain at each step is indicated on each bar.

Model 3 had final accuracy=85.7% and AUC=0.911, altogether accounting for 62% of the variance with again Treg (30%), naïve CD4+T-cells (29%), CD8 (12%) and RF (10%) contributing the most with less than 7% for all other variables.

Table 27 Modelling for the predicting of overall progression combining clinical + LS

	Logistic regression OR (95% CI) p-value (Wald test) n=210	
	unadjusted	**Model-3**
Smokers (ever)	2.520 (1.390-4.571) 0.002 (9.3)	3.158 (1.264-7.891) 0.014 (6.1)
HLA-SE positive	2.963 (1.651- 5.316) <0.0001 (13.3)	2.871 (1.212-6.800) 0.017 (5.7)
RF positive	3.710 (2.087-6.593) <0.0001 (19.9)	3.890 (1.681-9.004) 0.002 (10.0)
TJC78	1.113 (0.992-1.249) 0.068 (3.3)	1.261 (1.062-1.497) 0.008 (7.0)
CD8 T-cells	0.952 (0.917-.988) 0.009 (6.8)	0.908 (0.859-0.959) <0.001 (12.0)
NK cells CD56^{bright}	0.485 (0.178-1.317) 0.156 (2.0)	0.143 (0.027-0.751) 0.022 (5.3)
Naïve CD4 cells $^{\$}$	0.956 (0.934-0.978) <0.0001 (14.8)	0.892 (0.855-0.930) <0.0001 (28.8)
Treg CD4 cells$^{\$}$	0.651 (0.554-0.765) <0.0001 (27.0)	0.489 (0.374-0.630) <0.0001 (30.4)
Regulatory B cells	1.164 (1.021-1.325) 0.021 (5.3)	1.205 (1.1001-1.451) 0.049 (3.9)
Plasmablasts	1.119 (0.873-1.432) 0.375 (0.8)	1.303 (0.979-1.735) 0.070 (3.3)
Accuracy (%)	not applicable	85.7%
AUROC (95%CI) p-value		0.911 (0.871-0.951) <0.0001
Sensitivity (%) (95%CI)		83.9 (75-91)
Specificity (%) (95%CI)		80.3 (72-87)
PPV (%) (95%CI)		77.2 (70-83)
NPV (%) (95%CI)		86.2 (80-91)
Nagelkerke R square		62%
Hosmer& Lemeshow		0.287
Bias-corrected Somers D_xy		0.458

NK, natural killer; NKT, natural killer-T; Treg, regulatory T-cells; IRC, inflammatory-related cells; Exp-, expanded, $^{\$}$ normalised frequency, OR, odd ratio; AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. CI, Confidence interval. NB; Table presented only 10 variables (4 clinical + 6 LS) retained in the model

I further applied internal validation to model-3 to correct for optimisms using a bootstrapping technique (500 permutations), resulting in an optimised Dxy-value=0.458.

The high PPV/NPV of model-3 (77%/86%) suggests that it is possible to predict individuals who are likely to progress while identifying those who have low risk and may be monitored less often (even discharged). In this cohort, using dichotomisation based on the individual patient's probability to progress calculated from the logistic regression for each patient using model 3, the cohort was categorised into 2 risk groups. Based on the 80% sensitivity, a cut-off point in the distribution of probability values set at $p < 0.0001$, separated a group of patients that could be deemed high-risk (65/210). All but 6 progressed (91% accurate). 145/210 were deemed to be low risk probability and only 34 progressed (Chi square, $p<0.0001$) . Alternatively, within 78/210 cases of the low-risk group (i.e. probability <20%) only 6 (7.7%) progressed over 10 years.

This is further illustrated in **figure 52**, showing a time to progression comparing the high and low risk group using a survival curve analysis method to generate this plot.

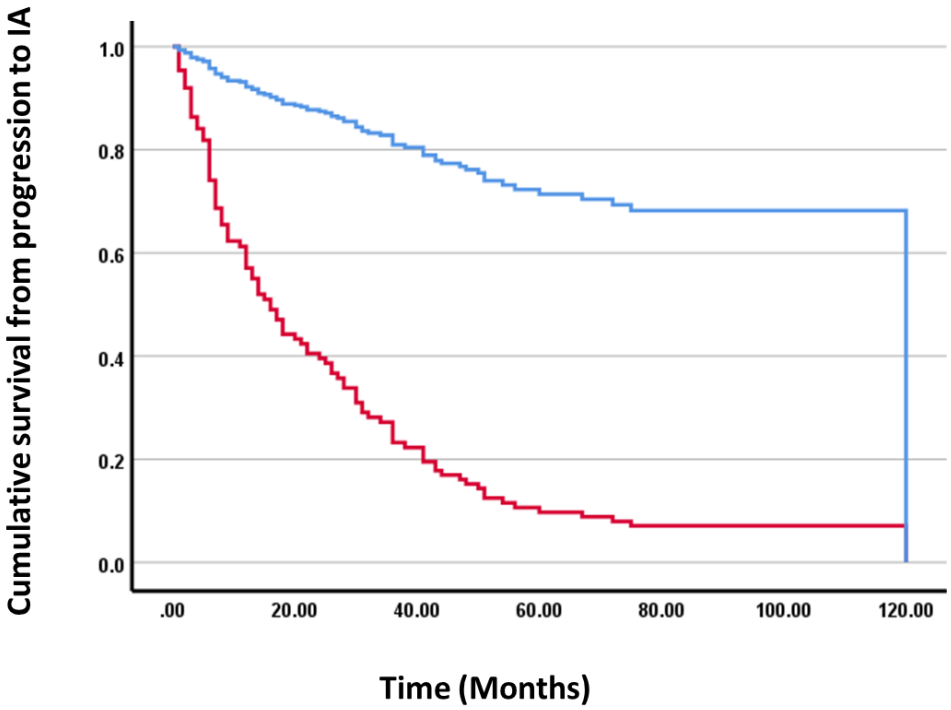


Figure 52 Time to progression based on classification using Model-3.
Survival plot analysis was performed after patients were dichotomised for High-risk (red line, n=65/210) and low risk (blue line, n=145/210) based on individual probability for progression calculated from the logistic regression using a cut-off at 80% sensitivity.

Summary of Model performance

The performance of the three models were compared using AUROC **(Figure 53).** The combination of both sets of data in Model-3 showed a +17% improvement in AUROC 0.911 (Cl95% 0.871-0.951) compared to Model-1 (AUROC 0.744, Cl95% 0.678-0.810). Comparing clinical data to LS alone (Model 2, AUROC= 0.862, 0.814-0.910) also increased accuracy (+5%)

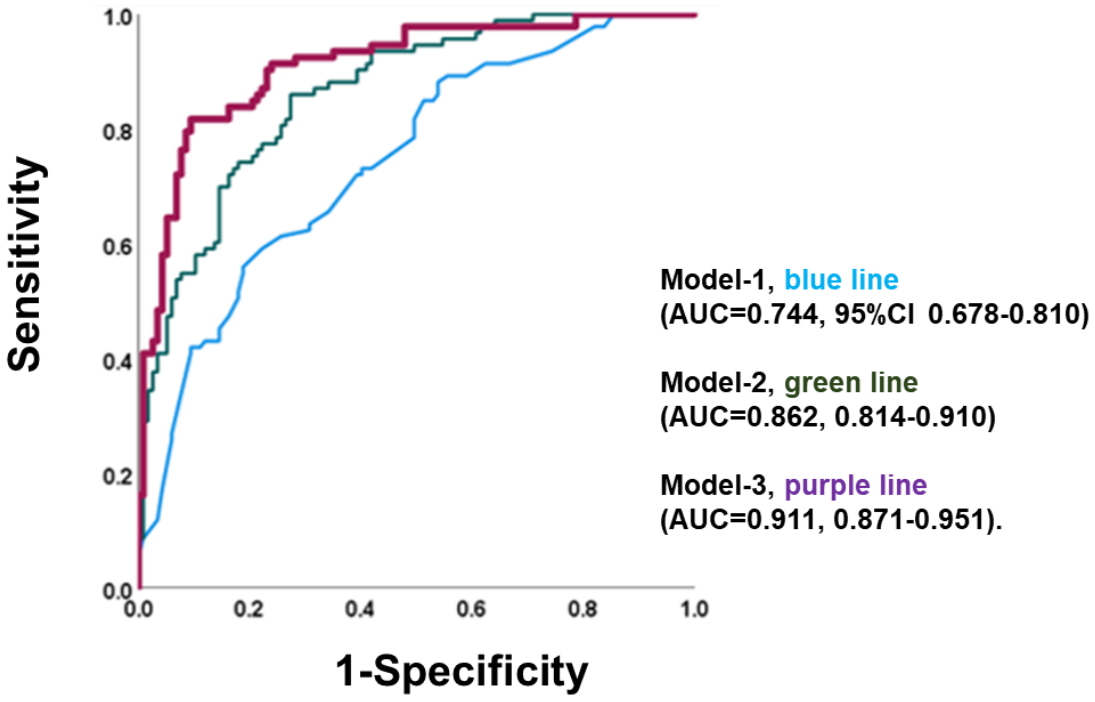


Figure 53 AUROC graphical representation of the logistic regression models.
Binary logistic regression models of the occurrence of progression to inflammatory arthritis (IA) were constructed using model-1 (Clinical data only) for 10 parameters (blue line), Model-2 (Flow- data only) for 18 subsets (green line) and Model-3 (Clinical+Flow data, purple line). Model-1 (AUC=0.744 95%CI 0.678-0.810) was inferior to Model-2 (AUC=0.862, 0.814-0.910) and Model-3 still showed added value (AUC=0.911, 0.871-0.951).

Modelling rapid progression to IA

Time to progression is widely distributed in this cohort ranging from 1-120 months. Different LS association may therefore be involved at different stages of the progression, and some may be more predictive of the onset of IA symptoms than others. 41/93 (44%) of the progressors did so rapidly and I re-analysed these progressors separately using a Cox regression.

Cox regression models were constructed using the same forward approach. Un-adjusted hazard ratio (HR) for time to progression were significant for 4 clinical variables (smokers, HLA-SE, RF, EMS, with CRP showing a trend, **table 28)**. Six LS (Treg, naïve CD4+T-cells, CD4-IRC, CD8+T-cells, NK-CD56^{dim} and B-reg) showed similar potential value.

I constructed similar models with clinical only (Model-4), the LS only (Model-5), and then combined both (model-6) for time to progression. The results are illustrated further in **figure 54**.

Model 4, retained 3 variables sequentially, starting with RF, TJC78 and finally smokers (AUC=0.702).

Model-5 (LS only) used 4 steps and retained 4 variables (AUC=0.773).

Model-6 (combined datasets) used 7 steps and showed better improvement (AUC=0.794).

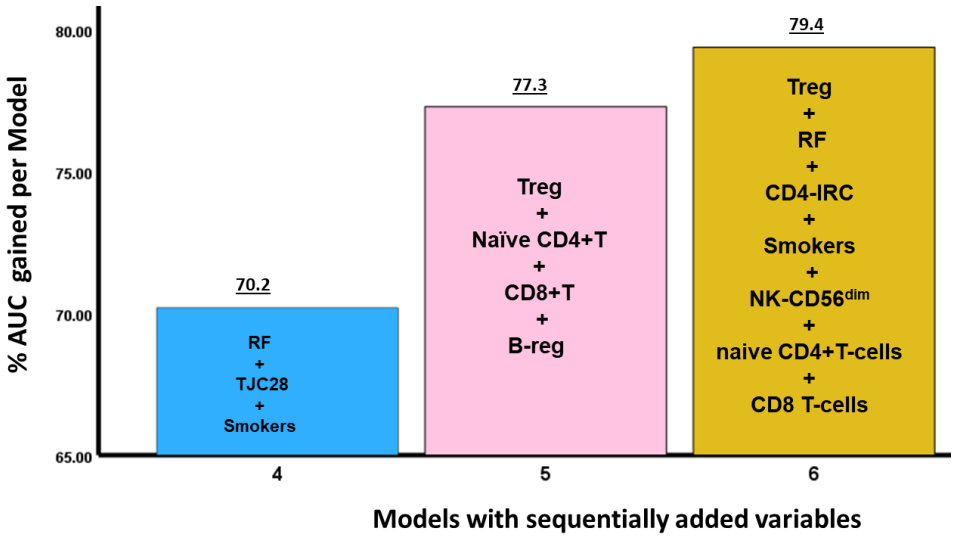


Figure 54 Cox regression model for rapid progression

The blue bar represents Model 4 (Clinical only) with the sequential 3 steps, purple bar represents Model 5 (LS only with 4 steps), Yellow bar represents Model 6 (Clinical + LS) showing 7 steps. Each model performance is indicated with AUC on each bar. Of note the use of accuracy as indicated for model performance is not applicable with Cox regression analysis. Hence, AUC was used.

Table 28 Modelling for the predicting of rapid (< 12 months) progression

	COX regression HR (95% CI) p-value (Wald test) n=158			
	unadjusted	**Model-4**	**Model-5**	**Model-6**
Smokers (ever)	2.005 (1.250-3.216) 0.004 (8.3)	2.662 (1.249-5.673) 0.011 (6.5)	not included in the model	3.688 (1.685-8.074) 0.001 (10.7)
RF positive	2.659 (1.743-4.056) <0.0001 (20.6)	4.767 (2.365-9.610) <0.0001 (18.9)		4.784 (2.173-10.532) <0.0001 (15.1)
TJC78	1.058 (0.984-1.137) 0.127 (2.3)	1.203 (1.070-1.352) 0.002 (9.5)		not retained
CD8 T-cells	0.960 (0.932-.989) 0.006 (7.4)	not included in the model	0.943 (0.905-.982) 0.005 (7.9)	0.950 (0.911-0.991) 0.018 (5.6)
NK cells CD56^{diml}	1.047 (1.005-1.090) 0.028 (4.8)		not retained	1.065 (0.999-1.136) 0.054 (3.7)
Naïve CD4 cells $	0.973 (0.958-0.989) <0.001 (11.1)		0.931 (0.906-0.956) <0.0001 (25.8)	0.954 (0.926-0.984) 0.003 (9.1)
IRC CD4 cells	1.055 (1.021-1.090) 0.001 (10.2)		not retained	1.113 (1.000-1.238) 0.049 (3.9)
Treg CD4 cells$	0.765 (0.689-0.851) <0.0001 (24.6)		0.638 (0.544-0.747) <0.0001 (30.8)	0.656 (0.554-0.776) <0.0001 (24.1)
Regulatory B cells	1.123 (1.024-1.232) 0.013 (6.1)		1.210 (1.053-1.390) 0.007 (7.1)	not retained
AUROC (95%CI) p-value	not applicable	0.702 (0.697 - 0.704) p<.0001	0.773 (0.756 - 0.760) p<.0001	0.794 (0.785-0.791) p<.0001
Nagelkerke R square		20%	32%	42%
Bias-corrected Somers D_{XY}		not performed		0.533

ESR, erythrocytes sedimentation rate; CRP, C-reactive protein; EMS, early morning stiffness HLA (SE), human leucocyte antigen (shared epitope); TJC, tender joint count; RF, rheumatoid factor; NK, natural killer; NKT, natural killer-T; Treg, regulatory T-cells; IRC, inflammatory-related cells; $^{\$}$normalised frequency, OR, odd ratio; HR, hazard ratio, AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. CI, Confidence interval. NB: Table presented variables retained in at least one of the models

Model 6 had an AUC=0.794, suggesting prediction although less well than that for the overall cohort (AUC=0.911). The assumption or proportional Hazard was verified for all variables in model 6 and result was satifactory with global Schoenfield test p=0.047 **(Figure 55)**. A significant discrimination index (X^2) was calculated for each included variable, ranging from 31.0 for Treg to 0.00 for NK-$CD56^{dim}$ cells and Treg was shown as the most discriminating biomarker for rapid progression in the mode **(Figure 56).**

The bootstrapping approach was repeated and showed that an optimism corrected Dxy=0.533.

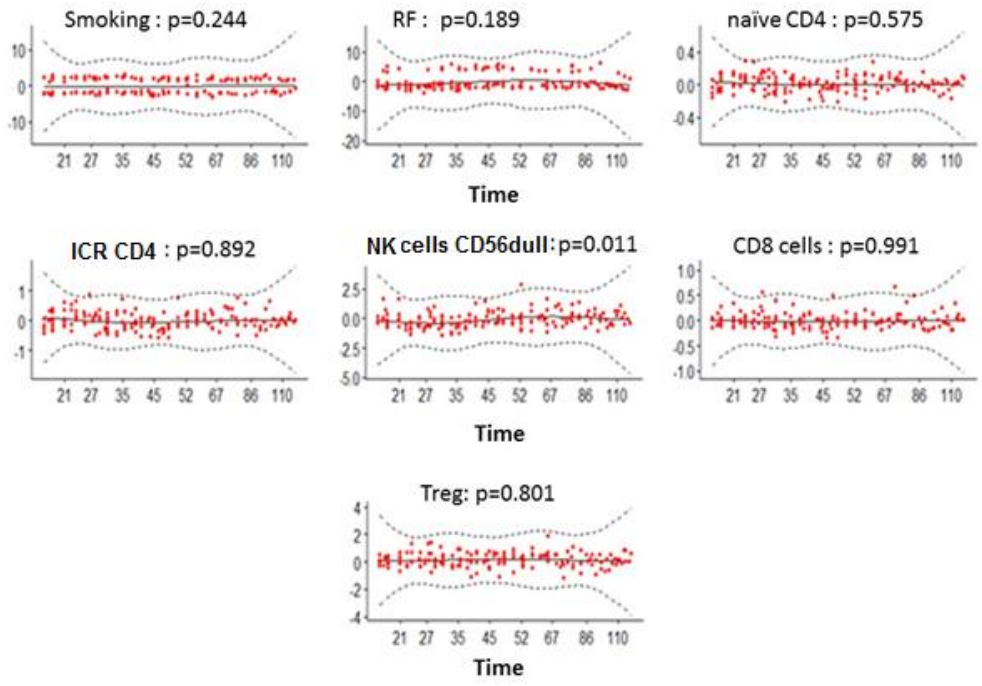


Figure 55 Testing the Proportional Hazards Assumption
Graphics show that the proportional Hazards Assumption was satifactory: global Schoenfield test P=0.047. The smooth curve was fairly level across the time horizon here, as opposed to substantially increasing or decreasing in level as time passes. So, proportional Hazards Assumption was satisfied. Individual Schoenfield test P-vlaues are indicated on each plot**.**

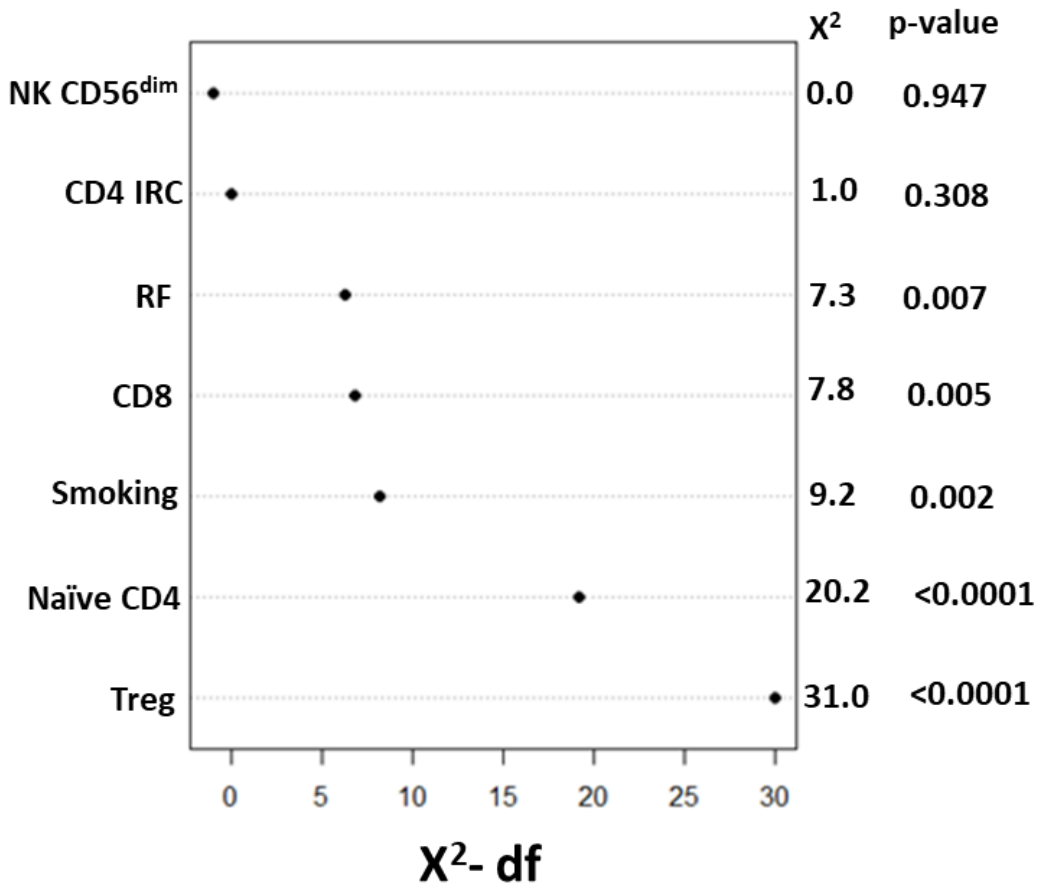


Figure 56 Variables contribution to Model-6
This showed the relative importance order of the predictors in the model with Treg as the most discriminating biomarker for rapid progression followed by naiveCD4, smoking, CD8, RF, IRC-CD4 and finally NK-CD56^{dim} cells.

Being able to identify individuals at high-risk of rapid progression (<12 months would allow for the design of clinical trials aiming at the prevention of progression within a short trial duration of only 12 months.

I therefore decided to use individual based dichotomisation of the risk of progression based on model 6 to categorise patients into high and low risk groups. In this rapid progression cohort (n=158), I used the individual X-beta score value > 2 calculated from the Cox regression for rapid progression to dichotomise patients for high and low risk. In this cohort, 22/158 participants

were dichotomised based on a high-hazard (individual X-beta score value > 2) for rapid progression and 20 (91%) indeed progressed as illustrated in **figure 57** using the time to progression curve plot.

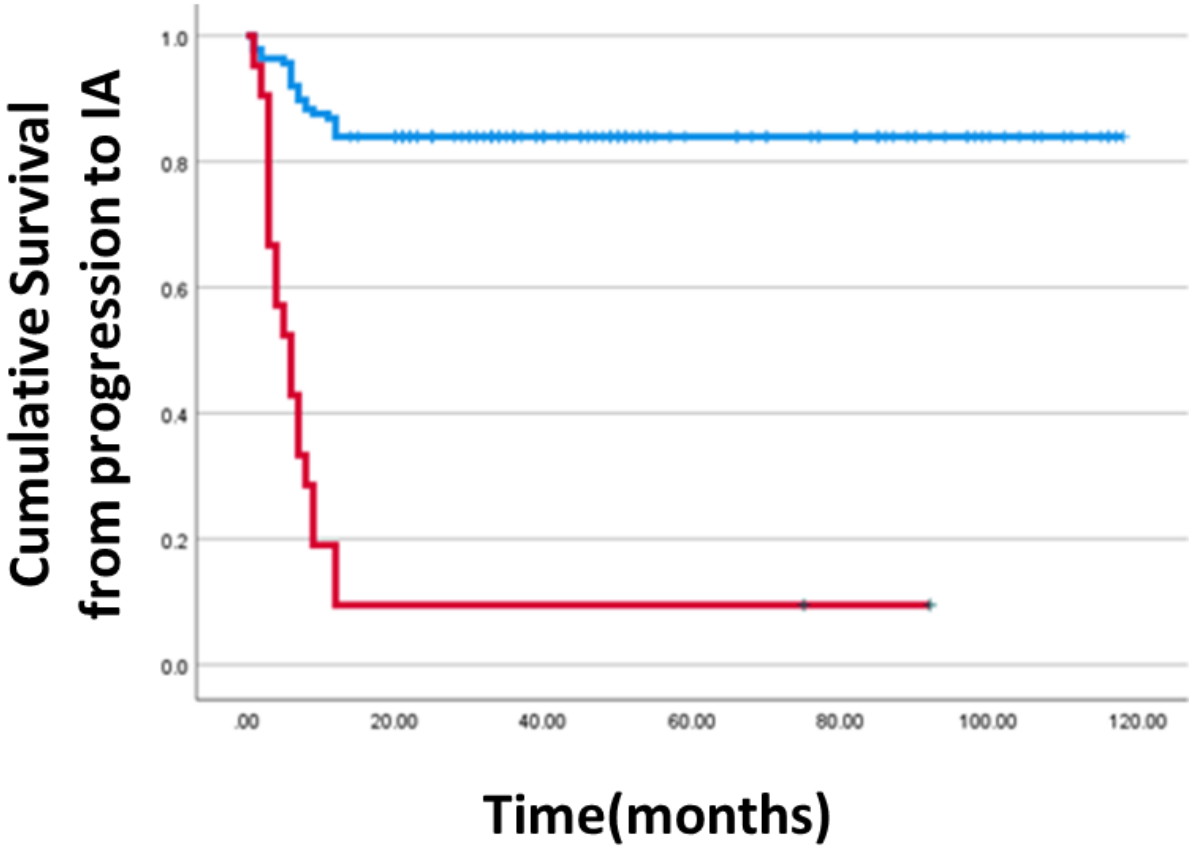


Figure 57 Time to progression based on the COX regression for rapid progression.
Survival plot analysis was performed after patients were dichotomised for High-risk (red line, n=22/158) and low risk (blue line, n=136/158) based on individual Hazard (>2) calculated from the COX regression.

Applicability for daily practice

Flow cytometry is routinely in clinical settings used in Leeds as well as worldwide. Therefore, I evaluated technically the gain in terms of accurate stratification of adding LS panels over using demographic/clinical data only.

I used first the demographic/clinical data (reference model) then sequentially added 1 to 4 FC panels starting with Treg (panel 1) and adding on panels 2, 3 and 4 in that order as shown in **(Table 29).** The model (logistics regression forward method) sequentially showed better accuracy 70 to 81.9% and AUC 0.744 to 0.911, demonstrating the value of added all 4 panels while selecting specifically some LS within each of them. **Figure 58 (upper panel)** presents the AUC of these 5 models and the gained obtained with each added panel.

The stratification showed 70% accuracy for the reference model. Adding the Treg panel increased accuracy to 75.2%. Adding the naïve CD4+T-cells panel showed further accuracy=80.4% while adding CD8+/NK-cells (lineage panel) and then B-reg (B-cell panel), only achieved a marginal improvement of accuracy to 81% and 81.9% respectively. The AUCs however were still improving with every incremental step and all 4 panels added value.

Individual participant's probability of progression obtained from the logistic regression analysis (from each of 5 Models**, Table 29**) was dichotomised (using a cut-off at 80% sensitivity which is deemed an acceptable risk clinically), setting cut-off value for each model to segregate high/low-risk groups.

Then a crosstabulation analysis was performed to determine the % of high-risk progressors and high-risk non-prog as well as low-risk progressors versus low-risk non-progressors **(Figure 58, lower panel)**.

Table 29 A practical approach to predicting overall progression to IA (n=210)

Logistic regression	Logistic regression				
	Model-ref	**Model + 1 panel**	**Model + 2 panels**	**model + 3 panels**	**model + 4 panels**
Clinical data	4 clinical variables	4 clinical variables	4 clinical variables	4 clinical variables	4 clinical variables
Flow panel -1 Treg		Treg CD4	Treg CD4	Treg CD4	Treg CD4
Flow panel -2 CD4 T cells			Naïve CD4	Naïve CD4	Naïve CD4
Flow panel -3 lineage				CD8 T-cells NK cells $CD56^{bright}$	CD8 T-cells NK cells $CD56^{bright}$
Flow panel -4 B-cell					B-reg PBs
Accuracy (95% CI)	70.0% (63.31-76.11)	75.24% (68.83 – 80.92)	80.0% (73.93 - 85.19)	81.0% (74.98 - 86.03)	**81.9%** (76.02 - 86.87)
AUROC (95%CI) p-value	0.744 (0.678-0.810) p<0.0001	0.827 (0.769-0.884) <0.0001	0.877 (0.832-0.923) <0.0001	0.898 (0.857-0.940) <0.0001	**0.911** (0.871-0.951) <0.0001
Dichotomisation Probability cut-off	high risk >0.510	high risk >0.500	high risk >0.430	high risk >0.400	high risk >0.390
Sensitivity (%) (95%CI)	55.9 (45-66)	68.8 (58- 78)	80.7 (71- 88)	81.7 (72- 89)	83.9 (75-91)
Specificity (%) (95%CI)	81.2 (73-88)	80.3 (72 - 87)	79.5 (71 - 86)	80.3 (72 - 87)	80.3 (72-87)
PPV (%) (95%CI)	70.3 (61-78)	73.5 (65- 80)	75.8 (68 - 82)	76.8 (69- 83)	77.2 (70-83)
NPV (%) (95%CI)	69.9 (64-75)	76.4 (70 - 82)	83.8 (77- 89)	84.7 (78 - 90)	86.2 (80-91)

AUROC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value.

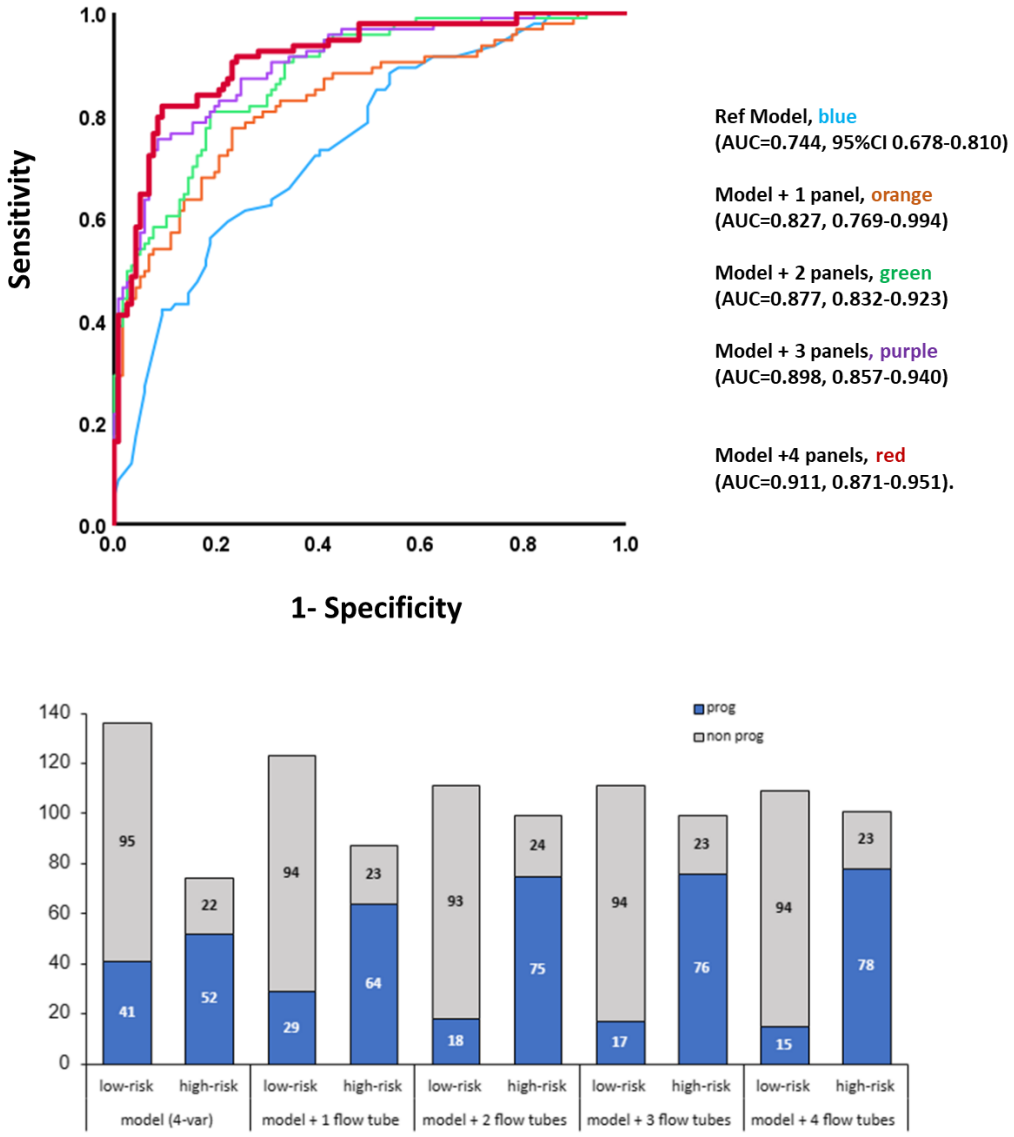


Figure 58 A. AUROC graphical representation of models for practical approach predicting overall progression to IA.

(Upper panel) Binary logistic regression models of the occurrence of progression to inflammatory arthritis (IA) were constructed in 5 logistic regression models including the demographic/clinical data only first and then, sequentially adding data from 1, 2, 3 and then 4 flow-cytometry panels. AUCs are indicated by the side **(lower panel), Overall performance of the prediction model using 1 to 4 flow cytometry panels**. Individual participant's probability for progression was dichotomised into high/low risk groups (based on 80% specificity) in 5 logistic regression models including the demographic/clinical data only first and then, sequentially adding data from 1, 2, 3 and then 4 flow-cytometry panels. Numbers of patients in both risk groups are displayed against the number of progressors (blue bars) and non-progressors (grey bar).

A similar analysis for imminent (rapid) progression suggested that only 3 panels would be needed while the performances clearly improved with the 3 steps (**Table 30).**

The stratification showed 17.9% accuracy for the reference model. Adding the Treg panel increased accuracy to 33.2%. Adding the naïve CD4+T-cells panel showed further accuracy=40% while adding CD8+/NK-cells (lineage panel) achieved a better improvement of accuracy to 45%. The AUCs however were still improving with every incremental step and all 4 panels added value from 64.8% to 79.1%

Table 30 Practical approach to predicting rapid progression to IA (n=210)

COX regression	**Model ref**	**Model + 1 panel**	**Model + 2 panels**	**model + 3 panels**
Clinical data	2 clinical Variables	2 clinical variables	2 clinical variables	2 clinical variable
Flow panel -1 Treg		Treg CD4	Treg CD4	Treg CD4
Flow panel -2 CD4/CD8			Naïve CD4 CD4-IRC	Naïve CD4 CD4-IRC
Flow panel -3 lineage				CD8 NK $CD56^{dim}$
Accuracy Nagelkerke's R^2	0.179	0.322	0.400	0.45
AUROC (95%CI) p-value	0.648 (0.644 0.652) <.001	0.735 (0.732 0.738) <.001	0.774 (0.771 0.777) <.001	0.791 (0.788 0.793) <.001

AUROC area under the roc curve. The 2 clinical variables included were RF and smoking

4.2.1.1.2 RA diagnosis phase (RA versus Non-RA)

Our group established the predictive value of the frequencies of circulating naïve CD4+T-cells but suggested no potential value of regulatory T-cells (Tregs), or IRC for RA diagnosis, as well as for the induction of remission at 1st treatment [5]. Here, I also investigated whether extended lymphocyte subsets phenotyping at the first visit in the early inflammatory arthritis clinical can provide an improvement over the performances of most recent model [5] that was using only the 3 CD4+T-cells subsets in combination with demographic and clinical data.

EAC Cohort description

Patients with early symptoms of inflammation in joints, were classified over 2 years of follow-up after referral as RA between 2008 and 2016, using the EULAR 2010 criteria (n=206) versus another form of arthritis (non-RA, n=100). These were disease-modifying anti-rheumatic drugs (DMARDs)-naïve patients, but the use of non-steroid analgesia and intra-muscular injection of steroid (depometrasone) within 3 months of the time of referral to the EAC, were permitted. A total of 306 patients with 4/5 LS panel available (see further detail **Figure 42, page 222)** could be included in this analysis using similar procedures as previously described for the at-risk cohort

Cohort outcome

Demographic/clinical data at baseline before evolving to RA or other forms of arthritis are described in **table 31.** Association between progression and data at inclusion suggested 6 highly significant parameters (RF, ACPA, CRP, TJC28, SJC28, or DAS, MWU p=0.0011 after correction) and another 2 parameters (Age, HLA-SE,) at trend (p=0.10).

This is consistent with the data previously reported by our group using similar EAC patients [5] as well as other similar cohorts [428, 735]

The individual calculated AUC suggests a predictive only value for 4 parameters, RF, ACPA, SJC28, or DAS (AUC>0.720, p<0.0001) and possible value for CRP and TJC28 (AUC>0.600, p<0.0001). The individual contribution to the prediction was however small for all non-significant parameters (<7%) but high for ACPA (62%) while lower for RF (37%) and < 20% for TJC28, SJC28, or DAS (Wald test).

Table 31 Association of demographic and baseline clinical data with outcomes of Early arthritis clinic (EAC) cohort (n=306)

	RA n=206(67.3%)	**Non-RA n=100(32.7%)**	adjusted p-value[&]	**AUROC (95%CI) p-value**	Unadjusted OR (95%CI) P value	**Wald**
Age (years)*	52.00 (43.75-62.25)	48.00 (33.25-58.00)	0.121	0.590 (0.518-0.661) 0.011	1.021 (1.004-1.039) 0.016	5.8
Gender [Female]	144(71.6%)	65(67.7%)	0.499	0.520 (0.449-0.590) 0.584	1.284 (0.769-2.142) 0.340	0.9
Smoking Never Ever	52(38.2%) 84(61.8%)	33(37.5%) 55(62.5%)	1.000	0.496 (0.419-0.574) 0.926	0.886 (0.540-1.454) 0.632	0.2
HLA-SE [positive]	23(88.5%)	14(56.0%)	0.143	0.662 (0.510-0.814) 0.047	0.941 (0.547-1.618) 0.826	0.1
RF [Positive]	93(60.4%)	14(14.9%)	***0.0011***	0.727 (0.664-0.791) <0.0001	6.170 (3.421-11.128) <0 .0001	36.6
ACPA[#] [Positive]	123(72.8%)	14(14.1%)	***0.0011***	0.793 (0.737-0.850) <0.0001	12.292(6.595-22.910) <0 .0001	62.4
CRP (mg/L) *	8.55 (0.00-24.50)	4.00 (0.00-11.20)	***0.0011***	0.636 (0.561-0.712) 0.001	1.013 (1.001-1.026) 0.039	4.2
TJC28*	9.00 (4.00-15.00	3.50 (1.00-10.00)	***0.0011***	0.691 (0.622-0.760) <0.0001	1.099 (1.053-1.147) <0.0001	18.8
SJC28*	5.00 (2.00-9.00)	1.00 (0.00-4.00)	***0.0011***	0.733 (0.670-0.795) <0.0001	1.273 (1.164-1.393) <0 .0001	27.7
DAS*	4.70 (3.52-5.61)	3.47 (2.32-4.30)	***0.0011***	0.730 (0.656-0.803) <0.0001	1.593 (1.293-1.964) <0 .0001	19.1
Duration of symptoms (weeks)*	24.00 (13.00-34.70)	28.00 (15.67-48.84)	0.330	0.422 (0.350-0.494) 0.030	0.985 (0.974-0.996) 0.009	6.8

Categorical data are presented as n (% of participant). * Numerical data are presented as median (Interquartile range values); p-values are indicated for Mann-Whitney and Chi-square tests for continues and categorical variables, respectively. ESR, erythrocytes sedimentation rate; CRP, C-reactive protein; EMS, early morning stiffness HLA (SE), human leucocyte antigen (shared epitope); S/TJC, Swollen/tender joint count; RF, rheumatoid factor; DAS, disease activity score, AUROC area under the roc curve. [&]tests adjusted for 11 comparisons (adjustment of the p-value was performed by applying the Bonferroni correction method for multiple comparison tests, 0.05/11 variables = 0.005 value or less considered significant after correction). CI, confidence interval. [#]268 patients had complete data.

I applied multivariate logistic regression with a forward approach to determine the predictive value of the demographic/clinical data alone **(Table 32).**

The model selected 4 parameters in a stepwise construction, starting with ACPA, and then sequentially adding SJC28, RF, and TJC28. **Figure 59** further showed the % accuracy gained on the addition of each step. The final model selected accurately predicted 83% of cases, with SEN and SPE=87% and 73% and a good PPV and NPV=86% and 73%, with an AUC=0.873 displayed later in comparison to other models on **figure 64 , page 283.** However, the model could only explain 48% of the variance for predicting progression to RA diagnosis (Nagelkerke R-square) and the individual variable contribution was highest for ACPA (41.3%), followed by SJC28 (10.2%) and less than 9% for the RF and TCJ28 (Wald score).

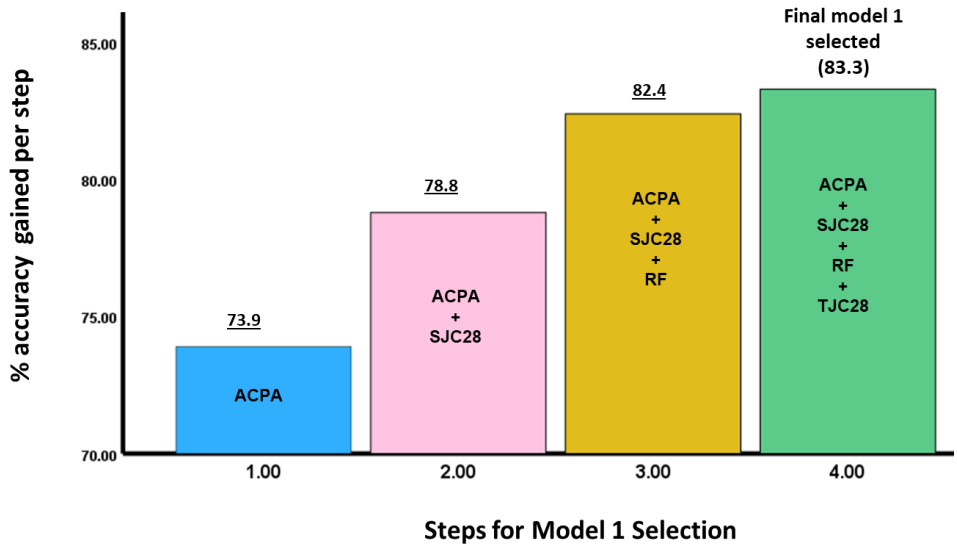


Figure 59 Accuracy gains in logistic regression modelling of clinical data : The blue bar represents step 1, the purple bar represents step 2, Yellow bar represents step 3 and green bar shows the final step retaining 4 variables. Model accuracy gain at each step is indicated on each bar.

Table 32 Unadjusted and multivariate modelling for predicting outcomes of Early arthritis clinic (EAC) cohort (n=306) using clinical variables only

Logistic regression	Logistic regression OR (95% CI) p-value (Wald test)	
	Unadjusted	**Model-1**
RF positive	6.170 (3.421-11.128) <0 .0001	2.862 (1.420-5.767) 0.003 (8.6)
ACPA [Positive]	12.292 (6.595-22.910) <0 .0001	10.004 (4.955-20.199) <0.0001 (41.3)
TJC28	1.099 (1.053-1.147) <0.0001	1.061 (1.001-1.125) 0.048 (3.9)
SJC28	1.273 (1.164-1.393) <0 .0001	1.205 (1.075-1.350) 0.001 (10.2)
Accuracy (%)	Not applicable	83.30 (77.61- 86.46)
AUROC (95%CI) p-value		0.873 (0.831-0.915) <0.0001
Sensitivity (%) (95%CI)		86.89 (81.51-91.18)
Specificity (%) (95%CI)		73.00 (63.20-81.39)
PPV (%) (95%CI)		86.89 (82.71-90.19)
NPV (%) (95%CI)		73.00 (65.10-79.67)
Nagelkerke R square		0.482
Hosmer& Lemeshow test		0.131

ESR, erythrocytes sedimentation rate; CRP, C-reactive protein; ; TJC, tender joint count; RF, rheumatoid factor; OR, Odd ratio; AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. NB: Only variables retained in the model were indicated in the table

Flow cytometry analysis of LS

FC analysis for the quantification of blood cell subsets was performed for lineage count, B-cell subsets; CD4+T-cell subsets; CD8+T-cell subsets; CD4+Treg. However, this was performed over 10 years (2010-2020) with all the caveats described in section **4.2.1.1 page 219**. As such, cohort had only 199 patients with at least 4/6 panel done presented a full dataset, and 109 patients were missing one or another panel. Randomness in the missingness was verified and was mainly due to no B-cell panel and no FoxP3 in Treg.

I imputed missing data using a similar process as to what I did for the at-risk cohort patients with >15/18 LS present. Furthermore, I imputed data for the CD4+T panel (due to issue with marker segregation) in 10% of patients who had complete data for all other LS. For the expanded memory CD8+T subset, I had an issue with poor separation between markers and difficulties in gating (as previously described) and I performed data cleaning to exclude this LS due to high missing datapoint (n=241), limiting the total number of LS used in further analysis in this cohort to 17. This allowed me to use 306 patients altogether.

Considering that only 4 subsets showed possible association before imputation and still showed association after imputation **(Table 34),** suggests that the multiple imputation procedure did not affect the dataset significantly (notably on those 4 subsets) when compared with the non-imputed dataset (**Table 33)**

Table 33 Data distribution characteristics of original dataset compared with imputed dataset for EAC cohort

Cell subsets	Original data Median (IQR)	Imputed data Median (IQR)	MWU p-value
CD4 T-cells	52.04 (46.06, 58.31)	52.62 (48.98, 56.14)	0.813
CD8 T-cells	18.97 (13.90, 24.01)	19.37 (16.23, 21.97)	0.579
B -cells	10.22 (8.26, 14.50)	11.00 (9.13, 13.05)	0.617
NK cells $CD56^{bright}$	0.39 (0.24, 0.57)	0.43 (0.20, 0.68)	0.666
NK cells $CD56^{dim}$	8.37 (4.90, 13.05)	9.21 (6.60, 11.70)	0.181
NKT cells	2.19 (0.97, 4.24)	2.87 (1.35, 4.51)	0.083
naive CD4 cells$	1.15 (-9.97, 11.43)	-1.33 (-10.77, 7.58)	0.975
memory CD4 cells$	-1.28 (-5.73, 3.03)	-1.59 (-4.84, 1.07)	0.442
IRCCD4 cells	2.10 (1.00, 5.80)	2.75 (1.16, 5.83)	0.415
Treg CD4 cells$	-1.27 (-2.73, 0.49)	-1.26 (-2.68, 0.07)	0.475
NaïveCD8 cells $	2.99 (-6.96, 11.56)	4.65 (0.55, 8.75)	0.704
MemoryCD8 cells$	-14.95 (-25.07, -7.77)	-14.84 (-18.53, -11.46)	0.987
IRC CD8 cells	13.50 (8.00, 23.40)	17.18 (13.27, 20.60)	0.086
naiveB-cells	0.38 (-14.13, 10.20)	-1.58 (-5.38, 3.82)	0.281
memoB cells	-2.84 (-11.71, 12.38)	-1.34 (-6.66, 3.83)	0.156
B-reg	4.00 (2.10, 5.70)	4.10 (2.69, 5.40)	0.509
PBs	0.70 (0.40, 1.30)	0.80 (0.32, 1.43)	0.953

Data are presented as median (1st and 3rd Interquartile values); CI Confidence interval; $ normalised subsets, NK natural killer, NKT natural killer-T, Treg regulatory T-cells, IRC, inflammatory-related cells; & MWU, Mann Witney U test . Breg= Regulatory B cells, PBs= Plasmablasts

The LS phenotyping results are presented in **figure 60** and **table 34**. In the lineage subsets (1st panel), none of the subsets showed a significant difference between RA and those progressing to other types of IA (non-RA) after correction for multiple testing (significant MWU $p<0.008$ after correction). Although, higher CD4+T % were individually associated with ($p=0.007$ before correction).

Frequencies of CD4+T-cell subsets (2^{nd} panel) were highly significantly associated with RA, showing reduced naïve ($p<0.0001$) and Treg subsets ($p=0.009$ after correction) but increased for IRC frequencies ($p=0.002$) which remained significant after correction ($p=0.006$). The memory subset did not show a difference between the 2 groups ($p=0.510$). CD8T-cell subsets (3^{rd} panel) and B-cell subsets (4^{th} panel) showed no significant association for any of the subsets between the RA and non-RA groups.

The overall results, therefore, still suggested that only the 3 CD4+T-cells were significantly association with progression to RA diagnosis (after correction) that nonetheless strengthens the concept of a major role for CD4+T-cells in predicting progression to RA.

Individual ORs were calculated, and the CD4 subsets had significant/possible predictive values for progression adding B-cells to the 2 CD4+ LS for CD4 naïve and Treg, and no longer for IRC, **(Table 34).** Importantly, only naïve CD4+T-cells showed a significant contribution to the prediction (Wald score 22%) while Treg contributed less (4.4%).

Despite not showing significant association by MWU test, only B-cells appears to have individual significant OR and contribute to the prediction (Wald test score, 7%). Hence, this is likely due to an overall effect of gating all lineage subsets as a 100% of the parental lymphocyte population.

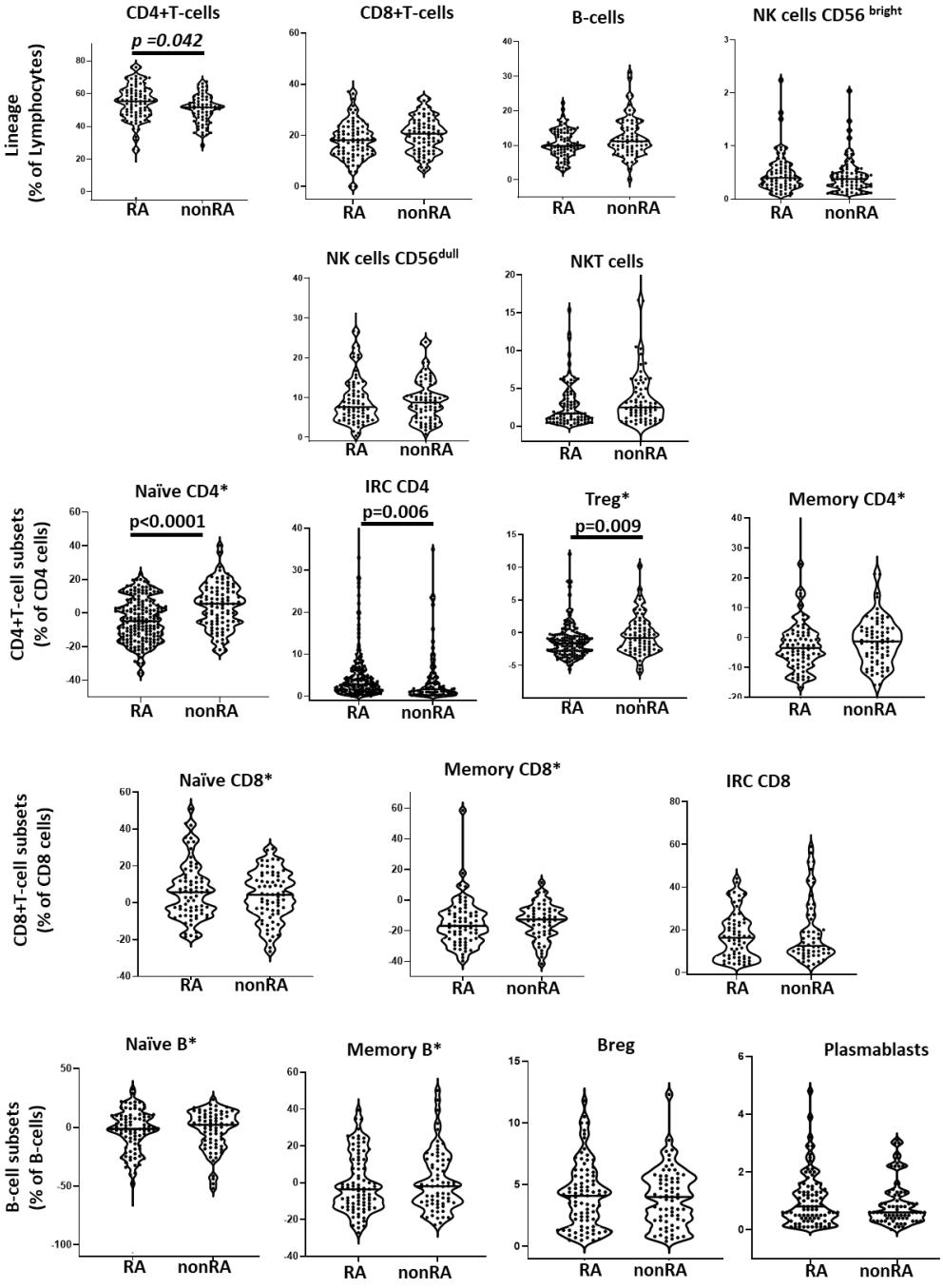


Figure 60 Frequency of lymphocytes lineage and subsets in early arthritis clinic (EAC) cohort, RA versus non-RA . LS were analysed by flow cytometry and data displayed as violin plots (each dot representing a patient) for Rheumatoid arthritis (RA, n=206) and non- Rheumatoid arthritis (Non-RA, n=100). Star (*) indicate LS that were normalised as previously described [5] P-value corrected for multiple testing (MWU test after correction) are indicated on the figure.

Table 34 Association of 18 blood cell subsets with outcomes of Early arthritis clinic (EAC) cohort (n=306)

	RA n=206(67.3%)	**Non-RA n=100(32.7%)**	MWU p-value[&]	**AUROC (95%CI) p-value**	Unadjusted OR (95%CI) P value	**Wald** test
CD4 T-cells	55.29 (46.93-61.88)	51.51 (44.40-56.24)	***0.042***	0.625 (0.539-0.712) 0.042	1.037 (1.005-1.071) 0.150	5.1
CD8 T-cells	18.15 (12.72-23.39)	20.61 (14.46-25.27)	0.282	0.408 (0.319-0.497) 0.282	0.972 (0.932-1.012) 0.167	1.9
B cells	9.85 (7.62-13.64)	11.17 (8.76-16.05)	0.270	0.408 (0.319-0.497) 0.270	0.912 (0.851-0.976) **0.048**	7.0
NK cells CD56^{bright}	0.40 (0.25-0.62)	0.38 (0.24-0.53)	0.443	0.537 (0.443-0.630) 0.443	2.721 (1.050-7.051) 0.234	4.2
NK cells CD56^{dull}	7.61 (4.90-13.05)	8.76 (4.92-13.26)	0.834	0.490 (0.396-0.584) 0.834	1.027 (0.976-1.081) 0.302	1.1
NKT cells	1.69 (0.87-4.10)	2.49 (1.23-5.05)	0.396	0.415 (0.326-0.504) 0.396	0.915 (0.837-1.002) 0.324	3.7
Naïve CD4 [$]	-4.71 (-13.81-4.75)	5.48 (-5.50-14.27)	***<0.0001***	0.328 (0.260-0.396) <0.0001	0.951 (0.931-.971) **<0 .0001**	21.8
Memory CD4[$]	-3.51 (-7.59-1.34)	-1.28 (-6.64-3.30)	*0.510*	0.435 (0.342-0.527) 0.510	0.998 (0.966-1.032) 0.920	0.0
IRC CD4	2.80 (1.30-6.00)	1.60 (0.60-4.50)	***0.006***	0.612 (0.540-0.691) 0.006	1.027 (0.982-1.073) 0.242	1.4
Treg CD4[$]	-1.64 (-2.88- -0.50)	-0.79 (-2.53-1.18)	***0.009***	0.394 (0.314-0.473) 0.009	0.897 (0.811-0.992) **0.035**	4.4
Naïve CD8 [$]	5.76 (-4.07-13.91)	3.78 (-6.71-14.75)	*0.419*	0.540 (0.443-0.637) 0.419	1.014 (0.988-1.039) 0.297	1.1
Memory CD8 [$]	-16.85 (-25.45- -7.77)	-12.64 (-21.85- -5.96)	0.276	0.441 (0.334-0.547) 0.276	0.993 (0.965-1.021) 0.596	0.3
IRC CD8	16.25 (7.45-23.70)	12.50 (9.00-26.50)	0.942	0.496 (0.391-0.602) 0.942	0.988 (0.960-1.017) 0.419	0.7
Naïve B [$]	-0.99 (-13.92-10.11)	2.28 (-14.00-11.00)	0.713	0.483 (0.389-0.575) 0.713	1.001 (0.981-1.021) 0.942	0.0
Memory B[$]	-2.595 (-11.29-13.24)	-2.009 (-11.29-13.73)	0.850	0.491 (0.399-0.583) 0.850	0.994 (0.974-1.015) 0.586	0.3
B-reg	4.10 (1.60-5.70)	4.00 (2.10-5.50)	0.844	0.509 (0.417-0.601) 0.844	1.021 (0.896-1.164) 0.756	0.1
PBs	0.80 (0.40-1.40)	0.60 (0.40-1.20)	0.460	0.535 (0.443-0.627) 0.461	1.244 (0.852-1.817) 0.258	1.3

Data are presented as median (Interquartile range values); CI Confidence interval; $ normalised subsets, NK natural killer, NKT natural killer-T, Treg regulatory T-cells, IRC, inflammatory-related cells; AUROC, area under the roc curve. [&] MWU, Mann Witney U test adjusted p-value for 17 comparisons (adjustment of the p-value was performed by applying Bonferroni correction method for multiple comparison tests for significance for each panel after correction [LC, 0.05/6=0.008; T-cell panels, 0.05/3=0.0166, B-panel=0.05/4=0.0125]

I then applied multivariate logistic regression with a forward approach to determine the predictive value of the LS alone (model-2) and of the combination of both (Clinical + LS) datasets (model-3).

Complete LS dataset after the imputation process as previously described for the at-risk cohort were used to construct the prediction models.

Model-2 initially used 7 steps. However, I chose the model achieved after step 3, based on a higher **Hosmer-Lemeshow** (HL) test =0.556) for that model retaining only 3 subsets (**Figure 61** and **Table 35**, naïve CD4, CD4, and Treg) compared to the ones using more subsets or less like step 1 or 2. Subsequent steps that added B, removing CD4, then adding on CD8, and finally NKT indeed reduced the goodness of fit (by 26% on average) with each addition.

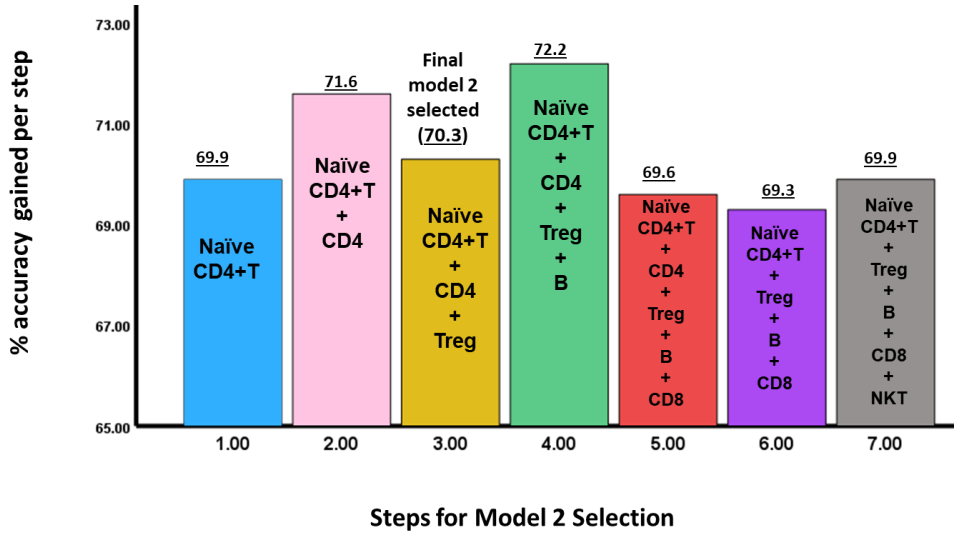


Figure 61 Accuracy gains in logistic regression modelling of LS only for RA:

The blue bar represents Step 1, the purple bar represents Step 2, Yellow bar represents Step 3 and green (Step 4), red (Step 5), purple bar (Step 6). Model accuracy gain at each step is indicated on each bar.

The model presented overall accuracy of 70.3% but accounted for only 18% of the variance with naïve CD4+T significantly contributing the most (Wald test, 28.5%) and other subsets <10%.

This model (Accuracy, 70.3%, AUC=72.7%) with 3 subsets in isolation was not better than model 1 (Accuracy, 82.4%, AUC=87.3%) suggesting that clinical and demographic characteristics are important to predict progression to RA.

I therefore proceeded to combine the 2 datasets (model 3).

Table 35 Unadjusted and multivariate modelling for predicting outcomes of Early arthritis clinic (EAC) cohort (n=306) using 17 LS only

Logistic regression	Logistic regression OR (95% CI) p-value (Wald test)	
	Unadjusted	**Model-2**
CD4 T-cells	1.037 (1.005-1.071) 0.025	1.055 (1.019-1.093) 0.003 (9.1)
Naïve CD4 [$]	0.951 (0.931-.971) <0 .0001	0.940 (0.918-0.961) < 0.0001 (28.5)
Treg CD4[$]	0.897 (0.811-0.992) 0.035	0.857 (0.769-0.955) 0.005 (7.7)
Accuracy (%)	Not applicable	70.26 (64.80 - 75.33)
AUROC (95%CI) p-value		0.727 (0.670-0.785) <0.0001
Sensitivity (%) (95%CI)		72.37 (66.47 -77.75)
Specificity (%) (95%CI)		59.18 (44.21 - 73.00)
PPV (%) (95%CI)		90.29 (86.81 - 92.93)
NPV (%) (95%CI)		29.00 (23.13 - 35.66)
Nagelkerke R square		0.182
Hosmer& Lemeshow test		0.556

[$] normalised frequency, OR, Odd ratio; AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. NB: Only variables retained in the model were indicated in the table

Modelling combining flow + Clinical data (Model 3)

Combining both datasets, Model-3 showed added value (**Table 36)**. Although 5 steps were suggested, a model selected based on retained 4 steps was best evidence of HL goodness of fit test=0.384. It selected ACPA, and SJC28, then naïve CD4+T (all p<0.001), and finally age (p=0.001). **Figure 62** further showed the % accuracy gained on the addition of each step. Model 3 did not show a great added value over model 1 in terms of accuracy (83.3% vs 84.6%), but a better AUC (87.3% vs 0.901, i.e., +3% added value) was observed. The fact that the model accounted for 53% variance compared to 48% for model 1 for predicting RA still suggest its relative superiority, thus naïve CD4+T contributed (21.3%) alongside ACPA (50.3%), SJC28(22.6%), and age (10.8%).

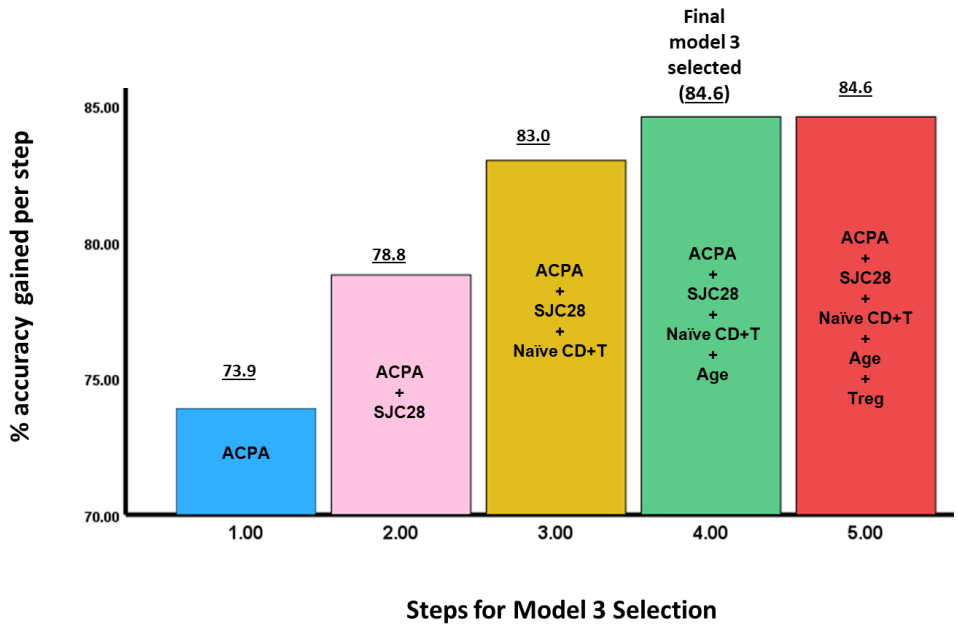


Figure 62 Accuracy gains in logistic regression modelling combining Clinical+ LS dataset for RA

The blue bar represents Step 1, the purple bar represents Step 2, Yellow bar represents Step 3 and green (Step 4), red (Step 5). Model accuracy gain at each step is indicated on each bar.

Our group previously reported value for the 3 CD4+T-cell subsets (naïve, IRC and Treg) in similar cohort [5] Modelling using a similar approach (Enter method as in this previous work) in these 306 patients combining the same 6 clinical and 3 CD4+T-cells variables, confirmed these previous data with a closely comparable accuracy (75.8% and AUC=0.821) in the original paper while IRC and Treg were not significant independent contributors to this model. Of note, in previous work by our group, ACPA and RF were excluded from the model as the most weighted variables used in the classification hence already accounted for in RA diagnosis.

Therefore, the use of a forward method that allow only independent contribution into the model suggest a better prediction (+ 8.8% accuracy and + 8% AUC).

Modelling flow + clinical data (model 4) in ACPA negative patients only

The most urgent need for a new classification biomarker remains in sero-negative RA, notably as this delays diagnosis by several months.

I, therefore, constructed another model 4 (**Table 36**, combining clinical + LS dataset) for 131/306 ACPA-negative patients in EAC cohort to identify set of potential predictive biomarkers. 46/131 ACPA-patients progressed to RA. Model 4 retained TJC28 ($p<0.0001$), naïve CD4 ($p<0.0001$), CD4 ($p=0.009$), duration of symptom ($p=0.035$), RF ($p=0.011$) and age ($p=0.051$) in that order using 6 steps **(Figure 63)** and predicted accurately 84% of ACPA negative patient who developed RA with both high sensitivity (82.50%) and specificity (85.71%). In the model, Naïve CD4+T contributed more (Wald test, 19.4%) followed by TJC28 (14.9%) and others between 3.8% - 6.9%.

In this cohort, I then segregated patients based on the predicted probability to progress calculated from the logistic regression in model 4 into 2 risk groups,

based on the 80% sensitivity cut-off point. A total of 20/131 could be deemed high-risk and all but 4 developed RA (80% accuracy) and 111/131 were deemed to be low risk but 27% (30/111) progressed to RA in this group (faulty prediction) In this particular group, patients achieved clinical classification for RA with a mean delay of 6.3 months. The non-RA group had a mean delay of 8.5 months as well, using LS would therefore allow 46 (16 high risk + 30 low risk) patients to be classified at baseline.

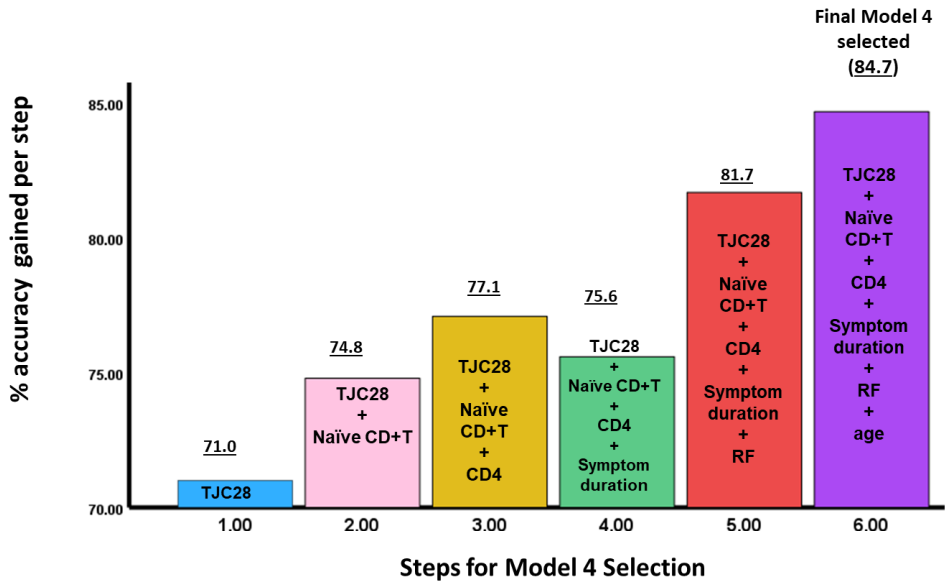


Figure 63 Accuracy gains in logistic regression modelling combining clinical+ LS dataset in ACPA negative patients:
The blue bar represents Step 1, the purple bar represents Step 2, Yellow bar represents Step 3 and green (Step 4), red (Step 5) and purple bar (step 6) with each showing the retained variables. Model accuracy gain at each step is indicated on each bar.

Table 36 Multivariate modelling for predicting outcomes of Early arthritis clinic (EAC) cohort (n=306) using the combined dataset

Logistic regression	Logistic regression OR (95% CI) p-value (Wald test)	
	Model-3	**Model 4 (ACPA negative patients) n=131**
Age (year)	1.040 (1.016-1.065) 0.001 (10.8)	1.035 (1.000-1.071) 0.051 (3.8)
RF positive	Not retained	5.298 (1.454-19.303) 0.011 (6.4)
ACPA [Positive]	12.875 (6.357-26.080) <0.0001 (50.3)	Not included
TJC28	not retained	1.146 (1.070-1.229) <0.0001 (14.9)
SJC28	1.301 (1.167-1.450) <0.0001 (22.6)	Not retained
Duration of symptoms (weeks)	Not retained	0.971 (0.945- 0.998) 0.035 (4.4)
CD4 T-cells	Not retained	1.110 (1.027-1.200) 0.009 (6.9)
Naïve CD4 $	0.936 (0.910-0.962) < 0.0001 (21.3)	0.895 (0.852-0.940) < 0.0001 (19.4)
Accuracy (%)	84.64 (80.10 - 88.49)	84.73 (77.41- 90.42)
AUROC (95%Cl) p-value	0.901 (0.864-0.939) <0.0001	0.877 (0.816-0.938) <0.0001
Sensitivity (%) (95%Cl)	87.32 (82.10 - 91.48)	82.50 (67.22 - 92.66)
Specificity (%) (95%Cl)	78.49 (68.76 - 86.34)	85.71 (76.81- 92.17)
PPV (%) (95%Cl)	87.39 (86.28 - 93.22)	71.74 (60.07- 81.07)
NPV (%) (95%Cl)	73.00 (65.17 - 79.62)	91.76 (84.98- 95.64)
Nagelkerke R square	0.528	0.51.8
Hosmer& Lemeshow test	0.384	0.419

TJC, tender joint count; RF, rheumatoid factor; $ normalised frequency, OR, Odd ratio; HR, AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. NB: Only variables retained in each model were indicated in the table

Summary of Model performance

The performance of the three models when compared using AUROC **(Figure 64).** This showed that combining clinical variable and LS in model 3 had +3% improved performance in AUC=0.901 compared to Model 1 (AUC=0.873) and was able to explain about 53% of variance in progression towards RA compared to 48% for model 1. In ACPA negative patients, Model 4 also showed good predictive value discriminating between those developing RA from non-RA notably compared to a model not including LS (Model 5).

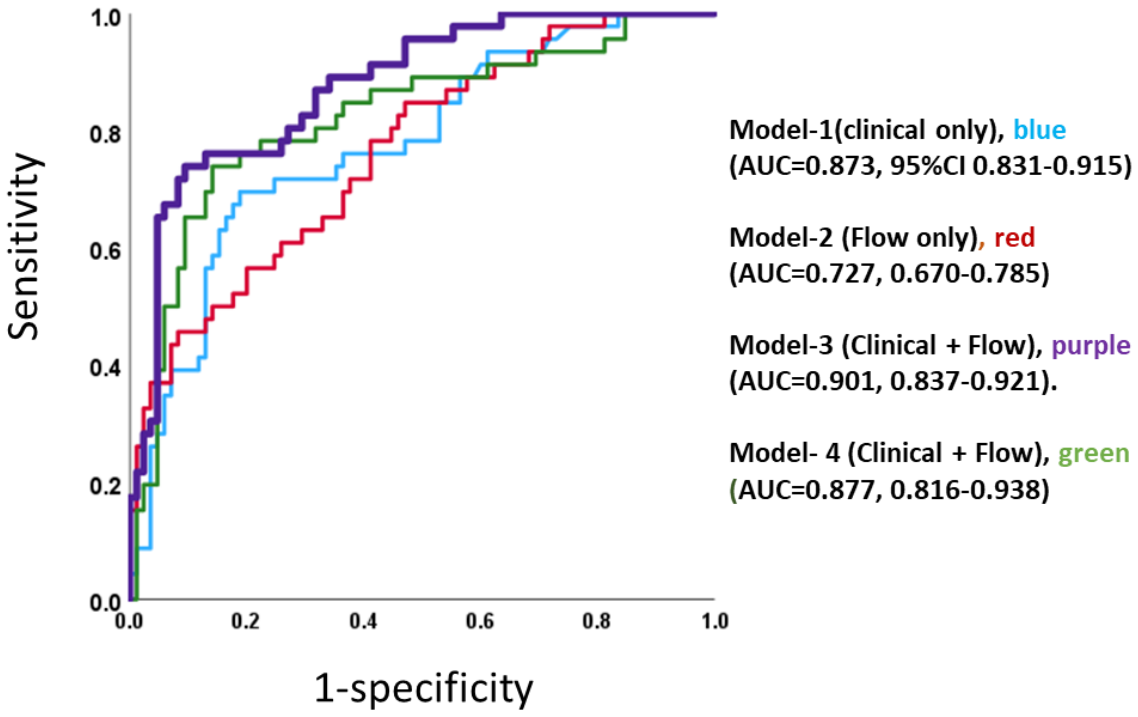


Figure 64 Performances of the models for RA.

The figure shows AUROC graphical representation of models. Binary logistic regression models of the occurrence of progression to RA diagnosis from inflammatory arthritis (IA) were constructed using model-1 (Clinical data only) for 10 parameters (blue line), Model-2 (Flow- data only) for 17 subsets (red line), Model-3 (Clinical + Flow data, purple line). Model-4 (green line) included Clinical + Flow data (ACPA negative patients only). AUC for each model indicated by the side.

4.2.1.1.3 MXT-treated RA cohort (MXT-induce remission versus non-remission)

To date, our group has proven that immunophenotyping of CD4+T cells in DMARDs-naïve early RA has value to predict clinical responses in MTX-treated drug-naïve RA patients. The contribution of other LS remains to be established. I hypothesized that dysregulations in circulating LC homeostasis at treatment initiation, could increase the accuracy of models (above the value associated so far with naïve CD4+T-cells). This would allow patients who are likely to respond to MTX to be discriminated from those who may not benefit from the use of monotherapy. Predicting MTX response at diagnosis will not only reduce cost of ineffective therapy but also help gain time in averting loss of function. Here I investigated the value of 17 LS to predict MTX-induced remission based on DAS<2.6 after 6 months of treatment.

Cohort selection characteristics and outcome

Participant selection was performed from the EAC patients with RA prescribed MTX at BL using same approach described for the at-risk cohort for availability of LS data added to the presence of clinical data allowing to establish response to MTX **(see Figure 42, page 222**). This cohort included 205 patients, of these achieving MTX-induced remission was seen in n=106 (51.7%) versus non-remission n=99 (48.3%).

The cohort comprised DMARD-naïve patients with early RA treated with a MTX initially at 15 mg/week using a Treat-to-target approach [736, 737] then escalated to higher doses (25 mg/week) over 8 weeks or the addition of other synthetics DMARDs, such as Sulfasalazine or Hydroxychloroquine, if remission was not achieved, at 3 months. Remission was evaluated at 6 months based on DAS<2.6

Demographic/clinical data at baseline are described in **table 37**. In univariate analysis I confirmed that only 3 parameters (CRP and DAS, MWU $p<0.030$, and smoking, x2 $p=0.060$ after correction) showed a trend for association with the achievement of remission. I observed no significant difference in the other 7 parameters; age, gender, duration of symptoms, RF, ACPA positivity, TJC28, SJC28 ($p>0.05$ after correction). Data, however, was not absolutely consistent with previous work (difference in OR ratio values and significance) [5] perhaps due to a smaller number of patients (n=70) in this previous work. For example, smoking was previously reported to have a highly significant association with remission, while it only showed a trend in the current study.

The individual contribution to the prediction was small for age and TJC (<5%) while >5% for the CRP, SJC28, smoking and DAS (Wald test).

Table 37 Association of demographic and baseline clinical data with MTX-treatment outcome at 6 months (n=205)

	Remission n=106(51.7%)	**Non-Remission n=99(48.3%)**	**MWU/ Chi-square p-value**	**AUROC (95%CI) p-value**	Unadjusted OR (95%CI) P value	**Wald test**
Age (years)*	59.00 (48.00-68.25)	55.00 (47.00-63.00)	*0.560*	0.577 (0.499-0.656) 0 .056	1.017 (0.998-1.037) 0.079	3.1
Gender [Female]	72(67.9%)	74(74.7%)	0.354	0.466 (0.387-0.545) 0.399	0.715 (0.389-1.316) 0.282	1.2
Duration of symptoms (weeks)*	23.50 (12.00-42.00)	26.00 (16.00-40.00)	0.233	0.451 (0.371-0.531) 0.234	0.996 (0.983-1.009) 0 .534	0.4
Smoking Never Ever	54(51.9%) 50(48.1%)	31(32.3%) 65(67.7%)	***0.060***	0.402 (0.323-0.480) 0.017	0.448 (0.248-0.785) 0.005	7.8
RF [Positive]	57(59.4%)	56(63.6%)	0.649	0.479 (0.395-0.562) 0.618	0.835 (0.460-1.515) 0.553	0.5
ACPA [Positive]	63(65.6%)	53(63.1%)	0.757	0.513 (0.428-0.597) 0.770	1.117 (0.606-2.058) 0.724	0.5
CRP (mg/L) *	7.7 (0.00-17.78)	13.400 (2.55-32.00)	***0.030***	0.380 (0.302- 0.458) 0.003	0.986 (0.975-0.996) 0.009	6.8
TJC28*	7.00 (3.00-13.75)	10.00 (4.00-17.00)	*0.310*	0.412 (0.331-0.491) 0.031	0.965 (0.930-1.002) 0 .061	3.5
SJC28*	4.00 (2.00-8.00)	6.00 (2.00-11.00)	*0.370*	0.415 (0.336-0.494) 0.037	0.939 (0.889-0.991) 0 .022	5.2
DAS*	4.10 (3.14-5.10)	4.79 (3.51-5.63)	***0.030***	0.378 (0.300-0.456) 0.003	0.744 (0.605-0.917) 0.005	7.7

Categorical data are presented as n (% of participant). * Numerical data are presented as median (Interquartile range values); p-values are indicated for Man-Whitney and Chi-square tests for continuous and categorical variables, respectively. ESR, erythrocytes sedimentation rate; CRP, C-reactive protein; EMS, early morning stiffness HLA (SE), human leucocyte antigen (shared epitope); S/TJC, Swollen/tender joint count; RF, rheumatoid factor; DAS, disease activity score, AUROC area under the roc curve. MWU Mann Witney U test adjusted p-value for 10 comparisons (adjustment of the p-value was performed by applying the Bonferroni correction method for multiple comparison tests, 0.05/10 variables = 0.005 value or less considered significant after correction)

As in the other cohorts, I performed the multivariate logistic regression with a forward approach to determine the predictive value of the demographic/clinical data alone (model-1**) (Table 38).**

Model-1 selected 3 parameters in a stepwise construction using 3 steps, starting with smoking then sequentially adding age, and then CRP (**Figure 65)**. This model accurately predicted 67.8% of patients achieving remission, with 48% sensitivity and 79% specificity, and PPV and NPV equals 71% and 58% respectively, with an AUC=0.714 **(Table 38).** Individual variable contributed not greater than 11% each. The model only explains 15% of the variance for predicting those who will achieve remission at 6months (Nagelkerke R-square)

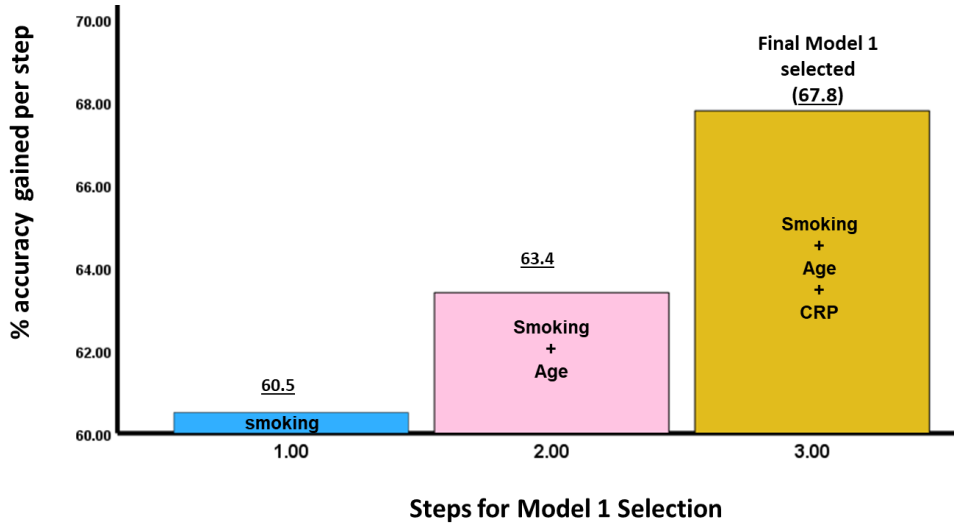


Figure 65 Accuracy gains in logistic regression modelling of clinical data for MTX-induced remission:

The blue bar represents step 1, the purple bar represents step 2, and Yellow bar represents step 3 showing the final step retaining 3 variables. Model % accuracy gain at each step is indicated on each bar.

Table 38 Unadjusted and multivariate modelling for predicting MTX-treatment outcome (n=205).

Logistic regression	Logistic regression OR (95% CI) p-value (Wald test)	
	unadjusted	**Model-1**
Age (year)	1.017 (0.998-1.037) 0.079	1.030 (1.008-1.052) 0.007 (7.2)
Smokers (ever)	0.448 (0.248-0.785) 0.005	0.364 (0.198-0.669) 0.001 (10.6)
CRP (mg/L)	0.986 (0.975-0.996) 0.009	0.982 (0.971-0.994) 0.003 (8.8)
Accuracy (%)	Not applicable	67.8 (56.42-70.01)
AUROC (95%CI) p-value		0.714 (0.643-0.784) < 0.0001
Sensitivity (%) (95%CI)		48.11 (38.30-58.03)
Specificity (%) (95%CI)		79.80 (70.54-87.20)
PPV (%) (95%CI)		71.83 (62.19-79.81)
NPV (%) (95%CI)		58.96 (53.84-63.89)
Nagelkerke R square		0.151
Hosmer& Lemeshow test		0.571

ESR, erythrocytes sedimentation rate; CRP, C-reactive protein; RF, rheumatoid factor; OR, Odds ratio; AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. NB: Only variables retained in each model were indicated in the table.

Flow cytometry analysis

For this cohort, I had the full 5 panel set for 74% of the patients while imputation was performed for 26% who were missing one or another subset mostly the B-cell panel and LC panel missing before 2015. Data cleaning was also carried out as previously described in the other 2 cohorts. Imputation process did not also impact the quality of the original dataset **(Table 39)**

Subsequent analyses were performed using the 17 LS in all available patients (n=205) after the data imputation process.

The data from univariate analysis for the 17 LS are displayed in **figure 66 and table 40.** In the lineage panel (1st panel), none of the lineage subsets showed association with achieving remission. For the CD4+T-cells subset (2nd panel), only the naïve subset at higher frequencies showed significant association with remission ($p<0.0001$) and remained significant after Bonferroni correction for multiple testing. Other CD4+T-cell subsets did not show any difference between the 2 groups ($p>0.05$). Similarly, for CD8T-cell and B cell subsets (3rd panel and 4th panel), no subset showed a significant difference between the 2 groups.

Table 39 Data distribution characteristics of original dataset compared with imputed dataset for the MXT-treated cohort

Cell subsets	Original data Median (IQR)	Imputed data Median (IQR)	MWU p-value
CD4 T- cells	52.33 (46.26, 58.68)	52.62 (46.48, 59.20)	0.891
CD8 T-cells	18.31 (12.83, 24.41)	18.57 (12.83, 25.28)	0.744
B-cells	10.75 (8.06, 13.63)	11.01 (8.13, 13.86)	0.712
NK cells $CD56^{bright}$	0.33 (0.20, 0.47)	0.35 (0.22, 0.51)	0.396
NK cells $CD56^{dim}$	9.10 (5.94, 13.63)	8.90 (5.90, 13.62)	0.779
NKT cells	2.24 (1.00, 5.13)	2.70 (1.06, 6.05)	0.216
naive CD4 cells$	3.95 (-8.00, 15.13)	4.20 (-8.03, 15.34)	0.947
IRCCD4 cells	1.60 (0.60, 3.70)	1.80 (0.65, 4.00)	0.846
memory CD4 cells$	-3.68 (-9.09, 1.77)	-4.07 (-10.02, 1.60)	0.817
Treg CD4 cells$	-1.36 (-2.59, -0.06)	-1.32 (-2.69, 0.09)	0.757
NaïveCD8 cells$	4.86 (-3.63, 12.42)	4.66 (-4.66, 12.32)	0.944
MemoryCD8 cells$	-17.33 (-25.89, -10.45)	-16.77 (-25.51, -8.62)	0.538
IRC CD8 cells	11.00 (6.00, 21.00)	13.08 (7.85, 24.10)	0.14
NaiveB cells $	-0.77 (-9.15, 8.18)	-1.22 (-11.70, 7.76)	0.653
MemoB cells$	-2.85 (-10.45, 6.04)	-2.17 (-10.24, 8.14)	0.634
B-reg	4.00 (2.35, 5.85)	4.10 (2.45, 6.08)	0.664
PB	0.70 (0.40, 1.30)	0.80 (0.40, 1.30)	0.389

Data are presented as median (1st and 3rd Interquartile values); CI Confidence interval; $ normalised subsets, NK natural killer, NKT natural killer-T, Treg regulatory T-cells, IRC, inflammatory-related cells; & MWU, Mann Witney U test . Breg= Regulatory B cells, PBs= Plasmablasts

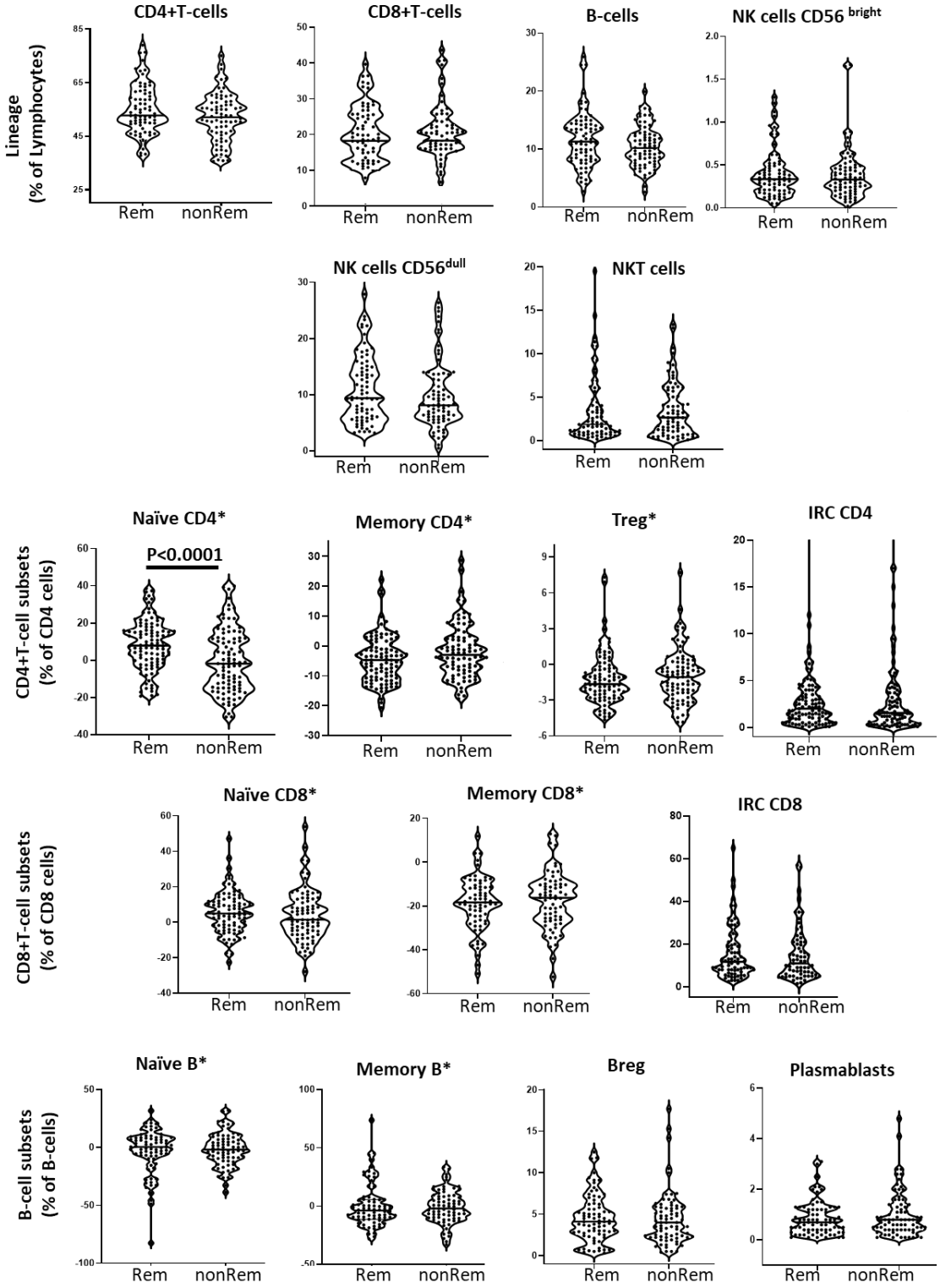


Figure 66 Frequency of lymphocytes lineage and subsets in MTX-treatment RA cohort.

LS were analysed by flow cytometry and data displayed as violin plots (each dot representing a patient) under remission (Rem, n=106) and non- remission (non-Rem, n=99). Star (*) indicate LS that were normalised as previously described (15). P-value corrected for multiple testing (MWU test) are indicated when significant.

Table 40 Association of 18 blood cell subsets with MTX-treatment outcome (n=205)

	Missing data	Remission n=106(51.7%)	Non-Remission n=99(48.3%)	&MWU p-value	AUROC (95%CI) p-value	Unadjusted OR (95%CI) P value	Wald test
CD4 T-cells		52.03 (45.37-57.58)	52.72 (46.95-61.13)	0.207	0.441 (0.349-0.532) 0.207	0.975 (0.944-1.007) 0.123	2.3
CD8 T-cells		18.26 (11.87-25.44)	18.31 (13.96-22.69)	0.478	0.466 (0.372-0.560) 0.478	0.989 (0.951-1.028) 0 .574	0.2
B cells		11.281 (8.140-14.276)	10.166 (7.742-2.540)	*0.108*	0.576 (0.484-0.668) 0.108	1.079 (0.995-1.169) 0 .065	4.0
NK cells CD56^{bright}		0.34 (0.21-0.49)	0.33 (0.20-0.47)	0.841	0.510 (0.416-0.603) 0.841	1.138 (0.360-3.596) 0.825	0.0
NK cells CD56^{dull}		9.39 (5.92-14.97)	8.11 (6.02-12.76)	0.248	0.555 (0.462-0.648) 0.248	1.034 (0.979-1.092) 0 .230	1.3
NKT cells	26%	1.89 (0.86-4.08)	2.68 (1.02-5.62)	0.486	0.467 (0.374-0.560) 0.486	0.991 (0.903-1.088) 0.854	0.0
Naïve CD4 $		7.91 (-0.63-16.15)	-1.88 (-14.75-11.83)	***<0.0001***	0.659 (0.582-0.736) <0.0001	1.038 (1.018-1.058) <0.0001	14.2
Memory CD4$		-4.75 (-10.57- 0.21)	-3.02 (-8.38-2.59)	*0.180*	0.421 (0.339-0.502) 0.060	0.964 (0.930-0.998) 0 .040	4.0
IRC CD4	0%	2.00 (0.90-4.00)	1.55 (0.50-3.85)	0.409	0.534 (0.453-0.616) 0.409	0.995 (0.940-1.052) 0.848	0.0
Treg CD4$	12%	-1.66 (-2.67- -0.31)	-1.06 (-2.71-0.14)	0.407	0.464 (0.380-0.549) 0.407	0.959 (0.838-1.098) 0.546	0.4
Naïve CD8 $		4.88 (-1.82-11.90)	1.43 (-6.60-12.42)	*0.420*	0.566 (0.478-0.653) 0.140	1.012 (0.989-1.036) 0.313	1.0
Memory CD8 $		-18.37 (-25.96- -11.21)	-16.27 (-26.27- -10.10)	0.608	0.474 (0.377-0.572) 0.606	0.992 (0.965-1.019) 0.542	0.3
IRC CD8	17%	11.75 (7.78-20.75)	11.00 (5.70-21.00)	0.565	0.528 (0.432-0.625) 0.565	1.005 (0.979-1.032) 0.707	0.1
Naïve B $		0.48 (-7.27-8.46)	-1.87 (-10.42-7.28)	0.595	0.525 (0.433-0.617) 0.595	0.996 (0.977-1.015) 0.690	0.1
Memory B$		-3.79 (-11.26-5.56)	-2.03 (-10.05-8.39)	0.510	0.469 (0.377-0.561) 0.510	1.004 (0.984-1.024) 0 .718	0.2
B-reg		4.10 (2.58-6.20)	4.00 (2.20-5.80)	0.584	0.526 (0.434-0.618) 0.584	1.011 (0.909-1.123) 0.845	0.0
PBs	26%	0.70 (0.38-1.20)	0.80 (0.40-1.40)	0.473	0.466 (0.375-0.558) 0.474	0.775 (0.512-1.171) 0.226	0.9

Data are presented as median (Interquartile range values); CI Confidence interval; $ normalised subsets; NK, natural killer; NKT, natural killer-T; Treg, regulatory T-cells; IRC, inflammatory-related cells; AUROC, area under the roc curve. & MWU, Mann Witney U test adjusted p-value for 17 comparisons (adjustment of the p-value was performed by applying Bonferroni correction method for multiple comparison tests for significance for each panel after correction [LC, 0.05/6=0.008; T-cell panels, 0.05/3=0.0166, B-panel=0.05/4=0.0125].

Using the same approach as before, I performed the modelling for the LS alone (model-2) using 17 LS, model excluded many subsets and retained only 2; 1st naïve CD4+T(p<0.0001), and then adding CD4+T-cells (p=0.004) with an overall accuracy=68.3%. Model-2 explains only 18% of the variance in predicting remission with naïve CD4+T contributing the highest (20%) and the other 2 each for <8% **(Table 41**).

Table 41 Unadjusted and multivariate modelling for predicting MTX-treatment outcome(n=205).

Logistic regression	**Logistic regression OR (95% CI) p-value (Wald test)**	
	unadjusted	**Model-2**
CD4 T-cells	0.975 (0.944-1.007) 0.123	0.948 (0.914-0.983) 0.004 (8.2)
Naïve CD4 $	1.038(1.018-1.058) <0.0001	1.052 (1.029-1.075) <0.0001 (19.9)
Accuracy (%)	Not applicable	68.30 (56.92-70.48)
AUROC (95%CI) p-value		0.721 (0.651-0.791) < 0.0001
Sensitivity (%) (95%CI)		49.06 (39.22-58.95)
Specificity (%) (95%CI)		79.80 (70.54-87.20)
PPV (%) (95%CI)		72.22 (62.68-80.10)
NPV (%) (95%CI)		59.40 (54.21-64.38)
Nagelkerke R square		0.183
Hosmer& Lemeshow test		0.194

$ normalised frequency, OR, Odd ratio; AUC area under the roc curve; PPV, positive predictive value; NPV, negative predictive value. NB: Only variables retained in each model were indicated in the table.

Combined modelling [clinical + LS data] (Model 3)

Next, combining both datasets (Clinical+ Flow), Model-3 predicted remission accurately in 70.2% of cases and showed added value (accuracy > model-2 and > Model-1) and AUC=0.734 **(Table 42**). The model included 3 variables starting with naïve CD4+T-cell (p<0.0001), CRP (p=0.002) and smoking (p=0.008) that showed significant association. **Figure 67** further presented the % accuracy gained on the addition of each step. The model confirmed the predictive value of naïve CD4+T-cell (p<0.001) contributing the highest (Wald test, 16.2%), and CRP (p=0.002, Wald test 9.2%) with smoking(p=0.008) making the least contributing (Wald test, 6.9%) but excluded DAS. Trying to include DAS instead of CRP made the model less fitting (Hosmer& Lemeshow test <0.0001). Data shown here are consistent with the findings previously reported by our group [5] and further validate the use of naïve CD4+T-cells (from 17 LS studied) as a predictive biomarker for MTX-induced remission in DMARD-naïve early RA patients. This suggests that a model using large cohort size with naïve CD4+T-cell datapoint would be sufficient to predict MTX-remission.

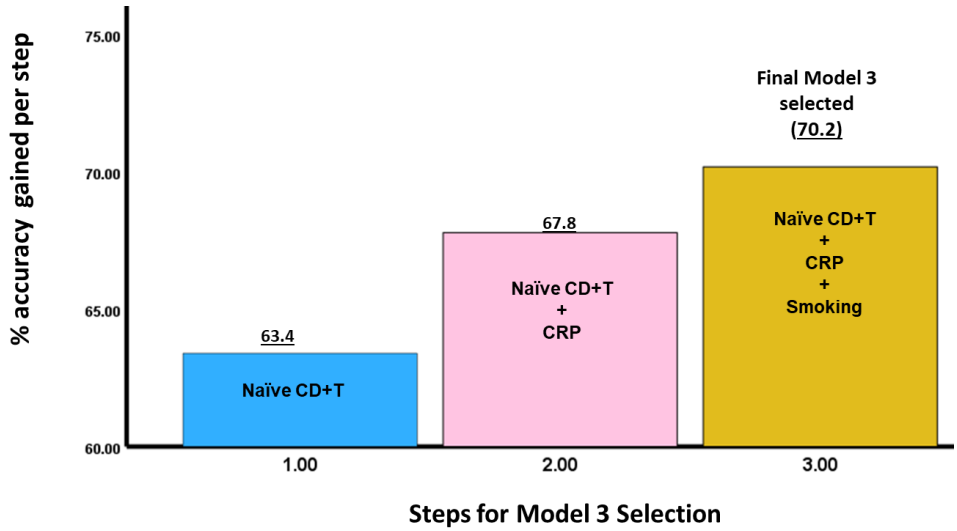


Figure 67 Accuracy gains in logistic regression modelling of Clinical +LS data for MTX-induced remission:
The blue bar represents step 1, the purple bar represents step 2, and Yellow bar represents step 3 showing the final step retaining 3 variables. Model % accuracy gain at each step is indicated on each bar.

With the help of 2 students, Helen Ng and Melanie Liu (who were allowed access to NHS computer over the pandemic while space access was restricted), I identified a few more patients (recruited from 2018 to 2019). This allowed me to perform a comparison of the early published model (n=120) [5] and all patients in my cohort with data for naïve CD4+T subset (n=277, 205 + 22 more patient identified) at the time of this analysis. I performed a second analysis with patients available (n=227) with panel 2 (naïve CD4+T-cell subsets irrespective of any other panel being done). Model 4, (clinical + naive CD4+T cell only) showed 71.9% accuracy and AUC=0.738. The model 3 showed lower performance over model 4 using naïve CD4+T alone. This confirms that the model using naïve CD4+T cell alone is indeed sufficient to predict MXT-remission while the smaller effect of the 22 more patients identified was eliminated as showing same data as if using only 205. Model performance is summarised and displayed in comparison to other models in **figure 68.**

Table 42 Unadjusted and multivariate modelling for predicting MTX-treatment outcome(n=205)

Logistic regression	Logistic regression OR (95% CI) p-value (Wald test)	
	Model-3 (n=205)	**Model 4 (n=227)**
Smokers (ever)	0.442 (0.241-0.810) 0.008 (6.9)	0.392 (0.22-0.699) 0.001 (10.0)
CRP (mg/L)	0.983 (0.971-0.994) 0.002 (9.2)	0.982 (0.972-0.993) <0.0001 (10.1)
Naïve CD4 $	1.043 (1.022-1.064) <0.0001 (16.2)	1.040 (1.021-1.059) <0.0001
Accuracy (%)	70.20 (58.42-71.86)	71.9 (63.63- 75.93)
AUROC (95%Cl) p-value	0.734 (0.665-0.804) < 0.0001	0.738 (0.673-0.804) <0.0001
Sensitivity (%) (95%Cl)	51.89 (41.97-61.70)	70.25 (61.26- 78.21)
Specificity (%) (95%Cl)	79.80 (70.54-87.20)	69.81 (60.13- 78.35)
PPV (%) (95%Cl)	73.33 (64.09-80.91)	72.65 (66.04- 78.39)
NPV (%) (95%Cl)	60.77 (55.39-65.90)	67.27 (60.34- 73.53)
Nagelkerke R square	0.208	0.217
Hosmer& Lemeshow test	0.104	0.842

CRP, C-reactive protein; $ normalised frequency, OR, Odd ratio; AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. NB: Only variables retained in each model were indicated in the table.

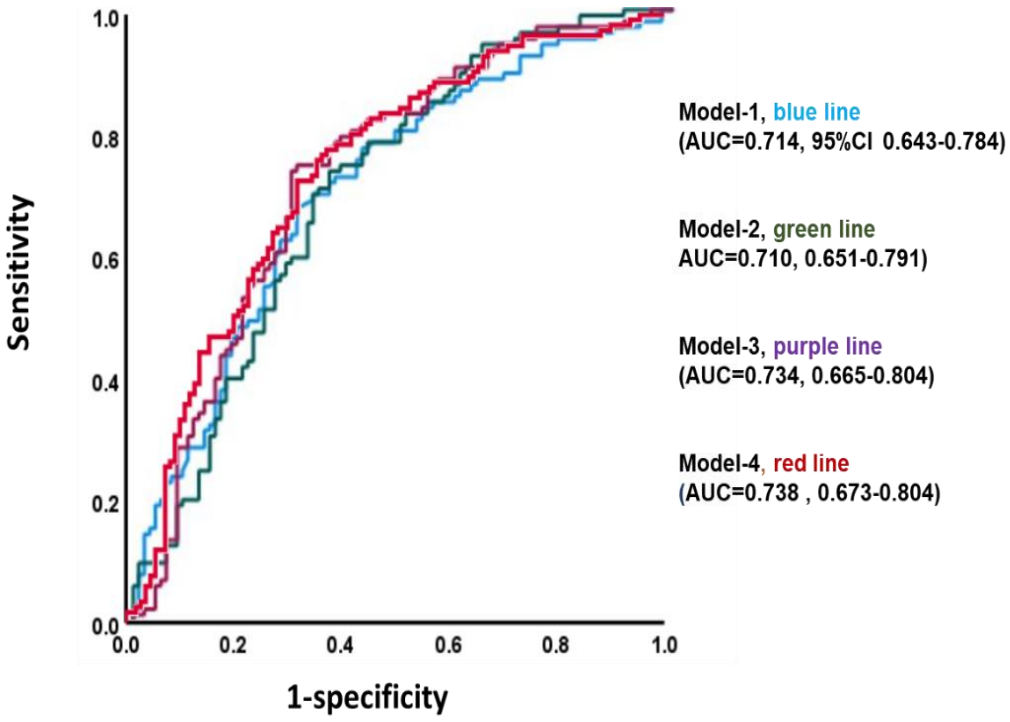


Figure 68 Performances of the models for MTX-induced remission

The figure shows AUROC graphical representation of models. Binary logistic regression models of achieving remission in MXT-treated RA were constructed using model-1 (Clinical data only) for 10 parameters (blue line), Model-2 (Flow- data only) for 17 subsets (green line), Model-3 (Clinical + Flow data, purple line) and, Moel-3 (AUC=0.734,0.665-0.804) added value (+2 accuracy). Model 4 (Clinical + naïve CD4+T) was comparable to Model 3 that included all the 17 subsets (red).

Comparison with model previously published from this cohort

I performed the forward logistics regression in the 2 groups for the clinical data + naïve CD4+T cells alone. For both datasets, the model selected the same 3 variables starting with naïve, CRP and finally smoking. The current model showed no improvement despite more patients **(Table 43 and Figure 69).** The main difference between groups was the patients from the control Clinical Trial who was recruiting all early RA at the time (VEDERA study) were included in the previous paper(62/120 patients with full dataset) while my cohort was mostly day-to-day patients included in a register and therefore reflecting more a real-life patient population of patients and missing clinical data at 6 months not allowing to select everyone. In addition, ACR 1957 criteria for classification were used in our previous work while EULAR 2010 are for more recent patients.

Table 43 multivariate modelling for predicting MTX-treatment comparing 2 groups using only 1 LS (naïve CD4+T)

Variables	Logistic regression OR (95% CI) p-value (Wald test)	
	Early cohort (2005-2010) n=120	**Current cohort (2011-2020) n=227**
Smokers (ever)	0.270 (0.108-0.672) 0.005	0.392(0.220-0.699) 0.001
CRP (mg/L)	0.974 (0.957-0.992) 0.004	0.982(0.972-0.993) <0.0001
Naïve CD4+ T cells	1.070 (1.038-1.103) <0.0001	1.040 (1.021-1.059) <0.0001
Accuracy (%)	78.26 (69.60- 85.41)	71.9 (63.63- 75.93)
AUROC (95%CI) p-value	0.816 (0.735-0.875) <0.0001	0.738 (0.673-0.804) <0.0001
Sensitivity (%) (95%CI)	78.33 (65.80- 87.93)	70.25 (61.26- 78.21)
Specificity (%) (95%CI)	78.18 (64.99- 88.19)	69.81 (60.13- 78.35)
PPV (%) (95%CI)	79.66 (70.01- 86.79)	72.65 (6.04- 78.39)
NPV (%) (95%CI)	76.79 (66.71- 84.52)	67.27 (60.34- 73.53)

CRP, C-reactive protein; $ normalised frequency, OR, Odd ratio; AUC area under the roc curve, PPV, positive predictive value; NPV, negative predictive value. NB: Only variables retained in each model were indicated in the table.

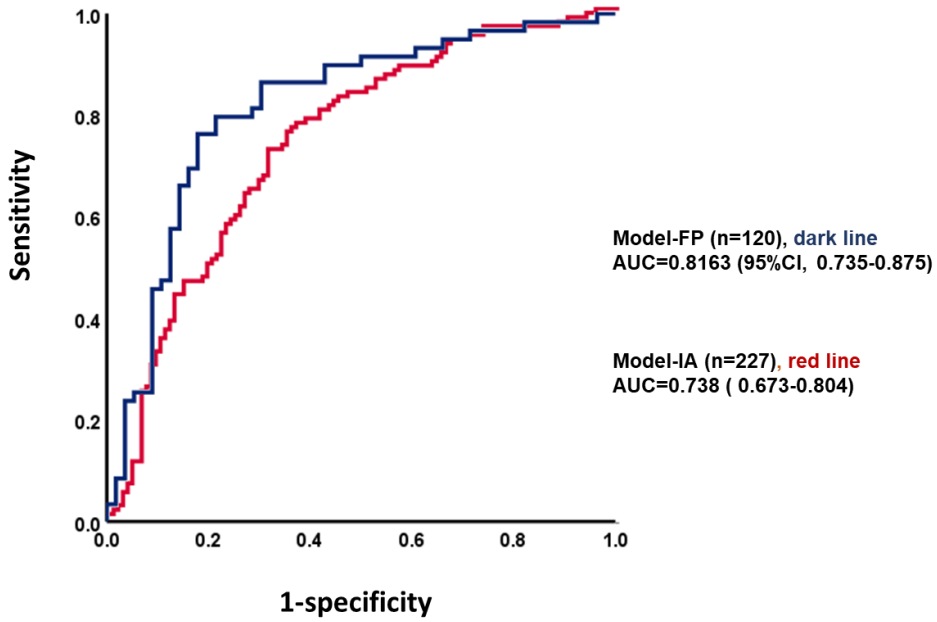


Figure 69 AUC Performances of the 2 models The figure shows AUC graphical representation of previous and current models.
The predicted probability scores from the binary logistic regression using (Clinical data + naïve CD4+T) were used to construct the AUROC graphs. Model FP (previous model, dark line) showed high performance index compared to Model-IA (Current model, red line). AUC for each graph is indicated**.**

Summary of blood cell subsets with predictive value across the stages of the inflammatory arthritis continuum (IAC).

Beyond the validation of the value of CD4+T subsets across IAC previously reported by our group, the TRIMID research group, the current study demonstrated added value of 3 novel cellular biomarkers (CD8, NK, and Breg) for predicting outcomes across the IAC **(Figure 70).**

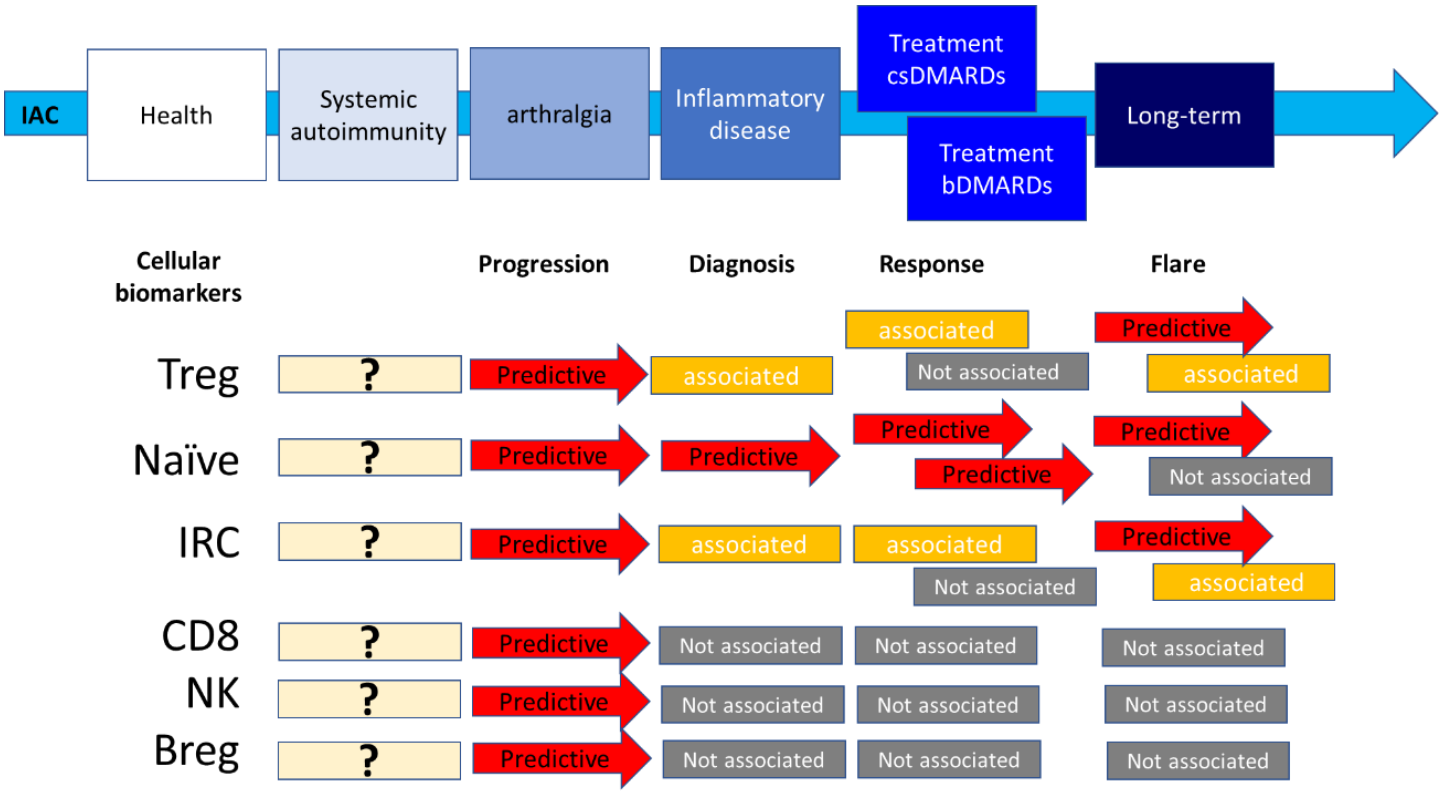


Figure 70 Summary of blood cell subsets with predictive value across each stage of the inflammatory arthritis continuum (IAC).
The diagram illustrated the biomarker value of 6 LS investigated across all phases of the IAC. Light yellow bars with question makers(?) indicate seropositive without arthralgia stage (with no identifiable biomarker yet). Red indicates subsets with predicate values. Yellow represents the subset with only association while gray bars are the subset with no significant association.

4.2.2 Validation of clinical utility of flow cytometry data using frozen samples

Introduction

For this aspect of my PhD, 36 cell subsets biomarkers across 4 LS panels (Lineage panel, T-cell panel, B-cell panel, and Th17 panel) were assessed at baseline (BL), 6 months, and 12 months in PsA patients from the GOLMePsA Trial. For the modelling analysis initially planned, I needed to get the full clinical dataset to assess the value of the flow data.

I, therefore, applied for access to the GOLMePsA clinical data including (1) treatment arms and (2) outcomes. Unfortunately, it was rejected by the MHRA, as the Trial was delayed due covid-19 pandemic. I could also only analyse 31 patients with full-time line (visits 1 to 5 at 12 months) due to time constraints and delays in reaching last visit. I also could only access the variable age (because I needed it to normalize the cell subsets) and gender. I still saw this as an opportunity to explore another type of statistics, in order to understand what could theoretically be used in terms of longitudinal data analysis. I choose to do a trajectory analysis based on time.

4.2.2.1 Data distribution

Data distribution for the 36 subsets for 31 patients at 3 time points (BL, 6 months, and 12 months) was visualised using a frequency distribution curve to access the pattern of distribution for each subset. Of the 36 subsets, only 8 were found to be normally distributed (passing the normality test, $p<0.05$). 2/36 subsets (naïve B-cell and HLA expression on B-cell) were particularly skewed to high value (right skewness). Of the 26/36 that showed left-skewness, 12 also had dual mode distribution (**Figure 71 C and D**). This skewness may suggest subsets different in groups of patients perhaps due to drug effects at different time-points particularly for subsets with bimodal distribution (D, left column) when compared to BL time-point (D, right column)

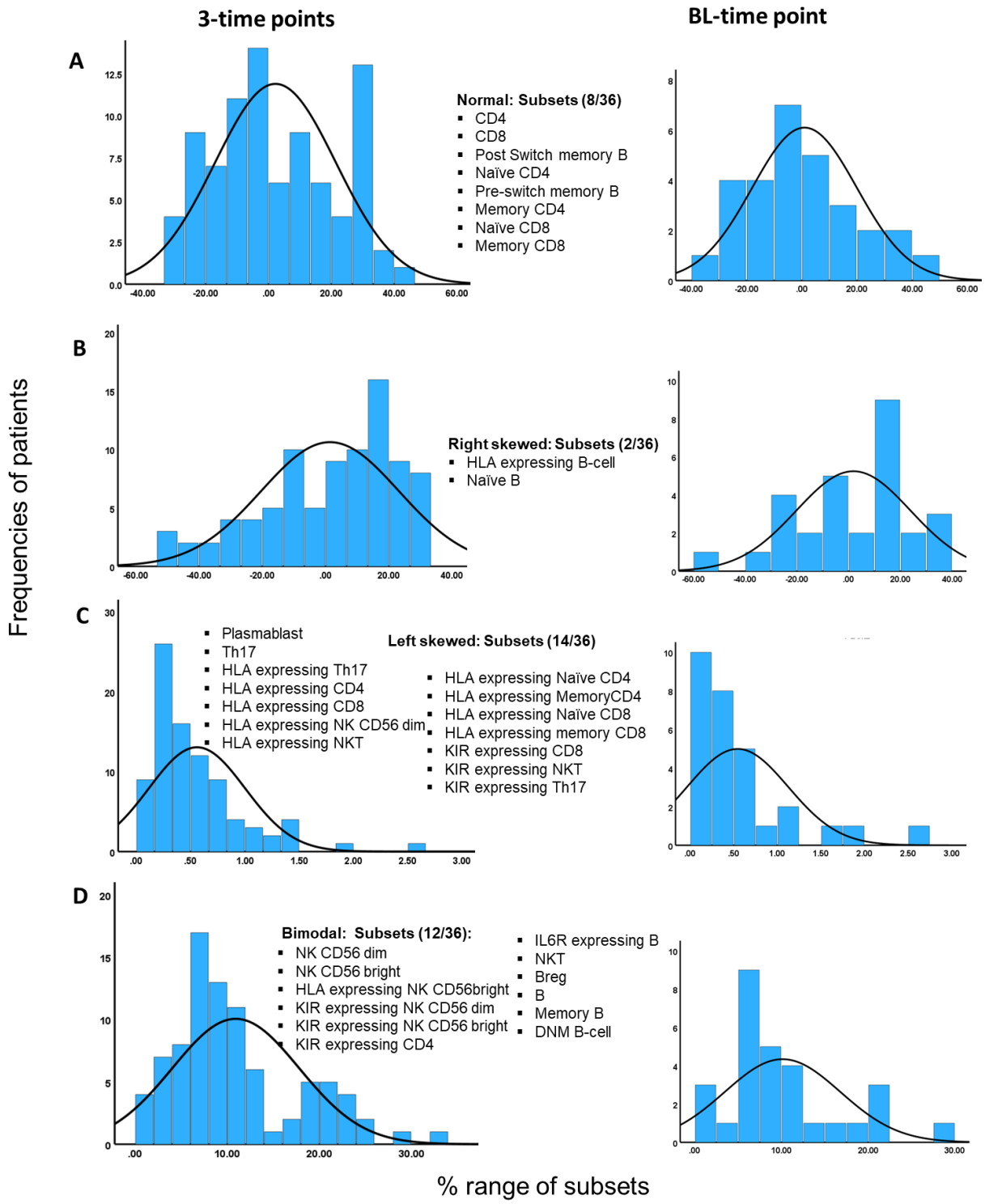


Figure 71 Representative data distribution curves of subsets(n=29) at 3 time points (left) and BL (right). (A) relatively normal distributed data with subsets indicated on the right (passed normality test, p>0.05), (B) right-skewed distribution failed normality test, p>0.05, and (C) left-skewed distribution, failed normality test, p>0.05) and (D) Bimodal distributed data (failed normality test, p>0.05) with subsets also indicated on the right.

4.2.2.2 Trajectory cluster analysis and association of subsets between clusters of patients

My objective was to identify hidden groups of patients (clusters) by considering the collective trajectories of subset frequency over time, which would provide insights into the distinctive features of each cluster[738] .To achieve this, I employed k-means clustering tailored for trajectory data analysis (Thanks to my supervisor Farag Shuweihdi supporting this analysis using the R package). A graph **Figure 72**) was used to visualise the quality criterion, employing the Calinski & Harabasz metric, specifically the Kryszczuk variant [739] for all previously discovered partitions. High values indicate partitions of "high quality," while low values suggest otherwise. Two clusters (labelled as "2") exhibit higher values of the Calinski & Harabasz metric suggesting clusters with informative insights compared to the other clusters, leading to two clusters being selected. The result presented in **Table 44** showed two clusters of patients (Cluster A, n=22 and Cluster B, n=7) which may speculatively suggest 2 subgroups of PsA patients with phenotypic lymphocyte differences.

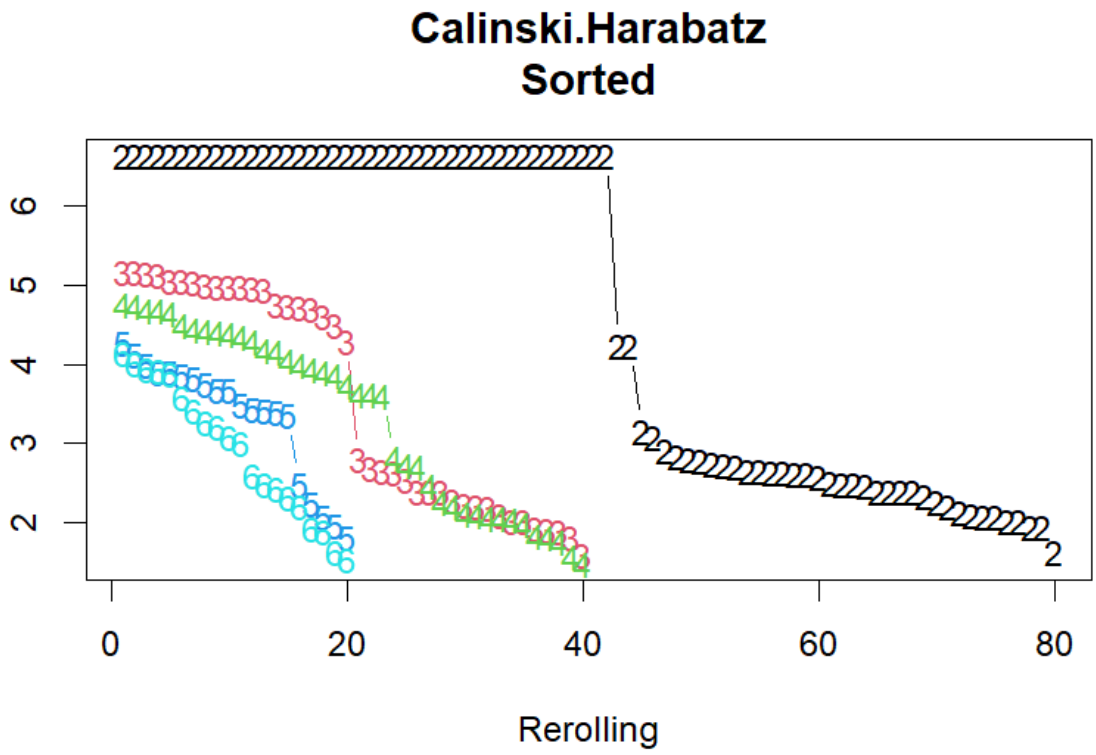


Figure 72 Estimation of the Number of Clusters Using Multiple Clustering Validity Indices. Suggested number of clusters are labelled as 2, 3, 4, 5 and 6.

Next, I performed a 2-way mixed ANOVA to determine the cell subsets that were driving/contributing to the clustering of patients over time. I observed that for a number of subsets (notably those not normally distributed), frequencies were significantly different between clusters A and B, at BL with no change over time remaining significant at 12 months but less so at 6 months. The frequencies of all other LS showed no significant difference between the 3 time points (BL, 6, and 12 months), suggesting that the patients clustering was driven by changes in subset frequency at BL rather than changes over time.

In more details, at baseline, memory B cells (p=0.0306 after correction) were significantly higher and DNM-B-cells lower (p<0.0001) in Cluster A compared to Cluster B. Naïve B-cells (p=0.0512) and Breg (p=0.0631) showed a trend for

difference between clusters as part of total B-cell population (100%) **(Table 44 and**

Figure 73). In contrast, all other subsets showed no difference at any time point, suggesting again no changes at BL and over time in these subsets.

At 12 months, I still observed the same difference in memory B cells (p=0.0108) and PreSM B cells (p=0.0180) and naïve B cells (p=0.0211), DNM-B cell (p<0.0001), and lower HLA expression on B-cell (p=0.0066) was also observed in cluster A patients. This analysis therefore suggested that there were 2 groups of patients in the study, mainly regulated based on B-cell subsets which was surprising in PsA as it may have been more expected on CD8+T cells being associated with HLA-B27.

Table 44 Association of 36 LS with cluster A versus B (n=29) at BL

BASELINE	A cluster (n=22) Median (IQR)	B cluster (n=7) Median(IQR)	Adjusted P-value
Age	36.25(29.50-59.00)	35.00(28.25-38.25)	0.372
Sex M/F n(%)	15(68.2%)/7(31.8%)	3(42.9%)/4(57.1%)	0.223
Panel 1			
CD4 (%)	39.65(9.25)	47.30(11.93)	0.6279
CD8 (%)	23.22(9.66)	26.04(10.59)	0.9858
B-cells (%)	11.28(7.72)	9.54 (5.18)	0.9930
NK CD56 bright (%)	0.87 (0.71)	0.39(0.43)	0.6930
NK CD56 dim (%)	11.24(6.45)	6.66(6.67)	0.6017
NKT (%)	8.64(5.71)	4.98(3.59)	0.6312
CD4 HLA (%)	2.03(0.98)	1.41(0.71)	0.6767
CD8 HLA (%)	2.29(1.65)	1.38(0.78)	0.7734
B HLA$_{(MFI)}$	2960.43(630.18)	3423.59(999.33)	0.8725
NK CD56 bright HLA%	5.52(3.35)	3.21(2.48)	0.3820
NK CD56 dim HLA%	0.52(0.38)	1.33(2.53)	0.2233
NKT HLA (%)	6.97(12.69)	5.06(4.64)	0.9992
CD4 KIR (%)	1.15(0.80)	1.31(1.29)	0.9983
CD8 KIR (%)	0.25(0.60)	0.10(0.07)	0.9108
NK CD56 bright KIR (%)	8.81(5.63)	10.67(6.84)	0.9948
NK CD56 dim KIR (%)	24.34(8.16)	25.47(10.31)	0.9999
NKTKIR%	16.40(12.51)	17.60(7.64)	0.9999
	Panel 2		
NaiveCD4 (%) *	1.24(19.81)	0.33(17.37)	0.9999
MemoryCD4 (%) *	-1.24(19.89)	-0.27(17.34)	0.9999
NaiveCD4 HLA$_{(MFI)}$	155.00(23.08)	191.49(109.84)	0.9999
memoryCD4HLA $_{(MFI)}$	157.33(12.66)	183.19(30.47)	0.7001
	Panel 3		
NaiveCD8 (%) *	22.77(16.11)	22.06(20.88)	0.9999
MemoryCD8 (%)*	-23.56(16.23)	-22.38(20.32)	0.9999
NaiveCD8 HLA$	159.21(39.81)	150.27(38.69)	0.9999
memoryCD8HLA$	161.47(33.19)	184.20(51.98)	0.9964
	Panel 4		
Naïve B (%) *	-4.10(21.84)	20.57(7.60)	0.0512
Memory B (%) *	1.41(17.92)	-20.64(6.23)	0.0306
Pre-SMB (%)	52.84(13.99)	37.69(10.87)	0.1049
Post-SMB(%)	39.50(12.47)	43.98(6.94)	0.9541
DNM (%)	7.67(3.65)	18.32(5.45)	**<0.0001**
B-reg (%)	3.05(1.44)	5.36(1.58)	**0.0631**
Plasmablaste(%)	0.57(0.36)	0.75(0.43)	0.9451
B IL6R$_{(MFI)}$	441.69(269.51)	593.64(431.84)	0.7437
	Panel 5		
CD4 Th17(%)	2.84(1.84)	2.17(2.10)	0.9189
CD4 Th17 HLA$_{(MFI)}$	87.12(11.81)	84.60(22.02)	0.9999
CD4 Th17 KIR$_{(MFI)}$	428.05(67.41)	524.27(200.10)	0.1459

* Normalized subset, $ levels of expression as median fluorescence index, adjusted p-value by 2-way ANOVA multiple comparisons test. DNM= double negative memory B-cells, SMB=switch memory B cells. All results are expressed as median (IQR) except for sex expressed in number (percentage).

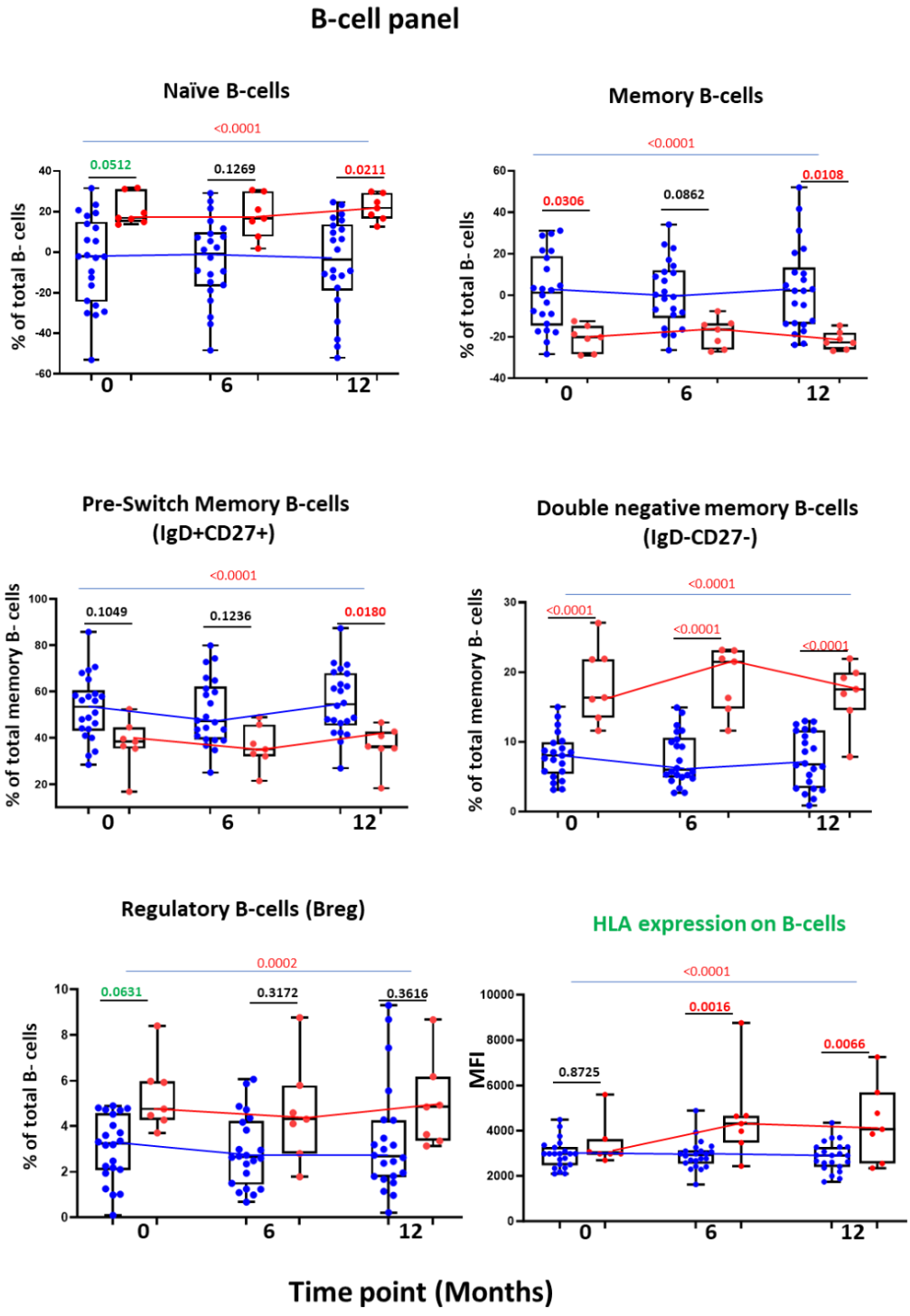


Figure 73 Frequency of B-cell subsets associated with change between Cluster A and Cluster B.

LS were analysed by flow cytometry and data was displayed as a box plot (each dot representing a patient) for Cluster A (blue colour, n=22) and Cluster B (red colour, n=7). Two way ANOVA P-values are indicated by the blue top line and p-value between groups

for the 3-time point indicated for trend (green), no significant difference (dark), and significant difference (red).

4.2.2.3 Unsupervised hierarchical clustering of the 36 subsets

To define groups of patients with similar LS profiles, unsupervised algorithm was used , that builds relationships between LS frequencies based on Spearman rank correlations. The algorithm was applied to log-transformed frequencies for 36 LS (n=30) at 3-time points (BL, 6M, 12M) and results are displayed as a heat map of data **(Figure 74).** The analysis segregated patients into 4 clusters based on 2 main groups of LS, suggesting different LS profiles. I observed that the clusters were driven by patients rather than time points as subset frequency from patients were closely related regardless of the time points and were therefor grouped together (the small box around the 3 time point for the same patients).

Altogether despite not being able to analyse data with respect to drug response this work suggests 2 types of PsA patients, based on a surprising B-cell signature, that is very stable over time which even treatment is hard to control disease.

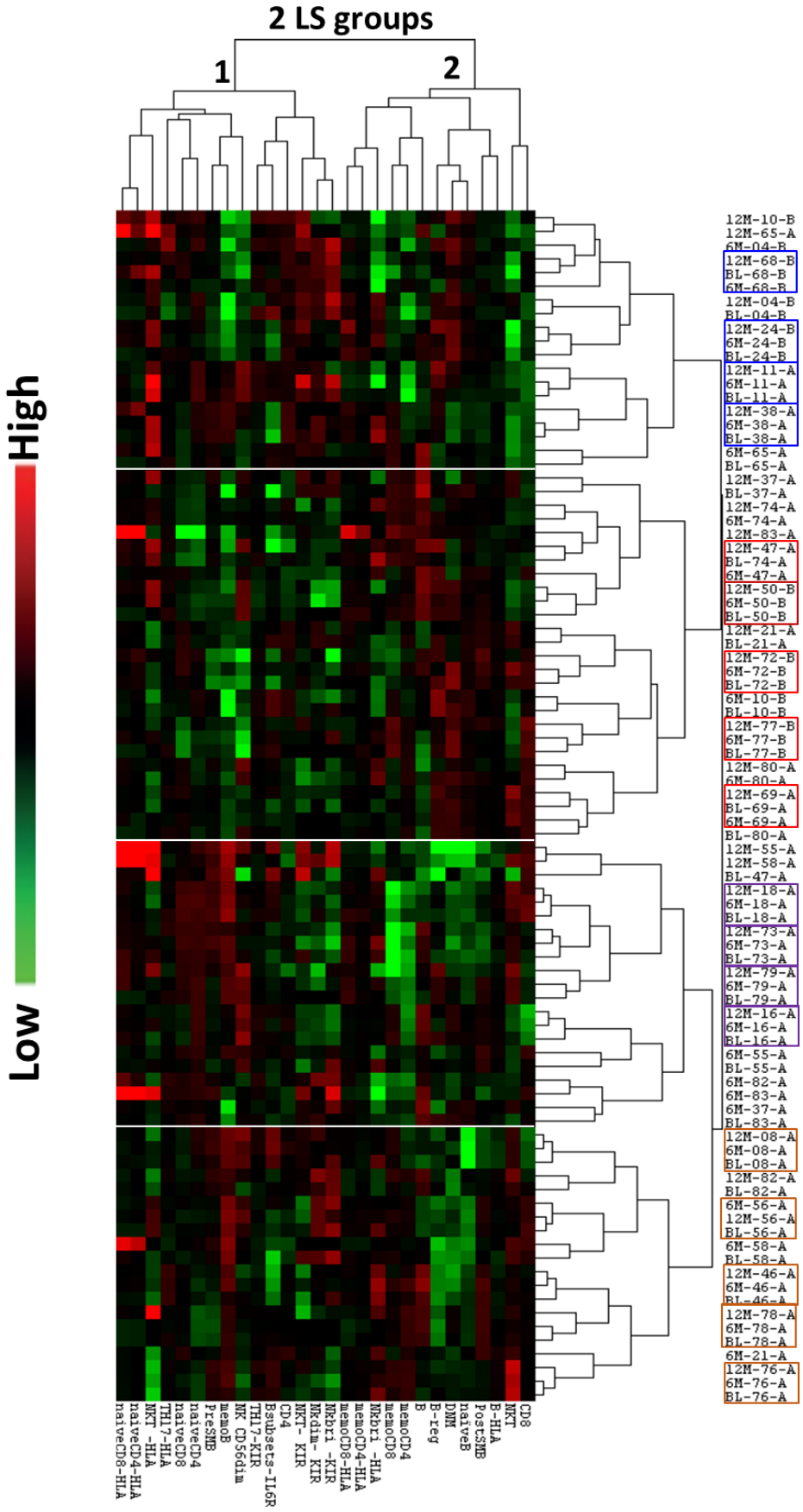


Figure 74 Unsupervised hierarchical clustering of the 8 subsets associated with cluster groups (n=29).

An unsupervised hierarchical clustering algorithm was applied to log-transformed frequencies for 36 LS and results are displayed as a heat map of data (red being the highest and green the lowest frequency observed for each LS). This clustering segregates patients into 4 clusters based on 2 main LS groups (upper dendrogram). 3 time point for the same patients grouped together in each cluster (small boxes).

Chapter 5 Discussion and Future Work

5.1 Technical Discussion

5.1.1 Validation of the use flow cytometry as a technology for biomarker research

A persistent challenge to accepting flow cytometry as a biomarker technology is the potential number of samples that cannot be analysed due to poor quality of blood, technical acquisition or data gating and/or reporting.

I performed a technical evaluation of the NHS flow cytometry data from the staining to data acquisition and analysis, using data already acquired by NHS immunology lab, to validate the extent of the impact of poor blood and faulty sample handling, gating error and data reporting by the NHS.

My results confirmed that four major problems arise. Faulty processing, poor instrument set-up, lack of compensation, while poor blood quality compromising the quality of a data may result from delays in reaching the lab. Consequently, this is limiting the amount of usable data when using the technology while still reported on NHS servers. My findings are consistent with the consensus report by a panel of flow cytometry experts from the American Thoracic Society that “poor handling and non-standardised sample processing is the step most likely to result in poor flow cytometry performance” [740].

Across the four panels analysed **(Figure 75),** human, involving particularly “missing antibodies” during staining, and technical errors presented the highest form of error encountered. It affected more of the Treg panel where FoxP3 was missing from the intracellular part of the protocol. This is a serious error as the definition of Treg is based on this marker more than any others. It led to reporting approximately 0.1%

Treg due to human error and could have led to major issues if this was a panel used for clinical decisions whereas it is still currently research as the TEEMS trial uses naïve cells for stratification and only records Treg for now. The errors due to missing CD127 are no less serious but here the point is that it should have been reported as a note for the clinician to make their decision fully informed as well as circulating the information about not being able to perform the test due to shortage of reagents. This suggests that a more careful training and QA/QC procedure needs to be in place for flow cytometry biomarkers to be taken at face value from any report.

Technical errors involving faulty acquisition either due to lack of compensation or arising from lack of adjustment in the instrument settings were more pronounced in B-cell panel analysis. This is mostly because this panel was introduced later and changed over time (from 8 to 5 markers). Nonetheless, reporting data from an uncompensated panel should have been spotted during a QA/QC process as it affects gating.

Poor blood sample quality may be attributable to the length of time between specimen collection and processing. This affected mostly the T-cell panel, most likely because the CD62L makers being more sensitive to experimental conditions [723, 724] are affected. Such delays in sample transit from clinics (in Chapel Allerton Hospital) to the NHS-lab (at St James hospital), via NHS-transport services (including change of vehicle at the Leeds General Infirmary) would need to be considered when optimising panels and reports to clinicians should account for such issues (or refuse the test based on the knowledge of its limitations).

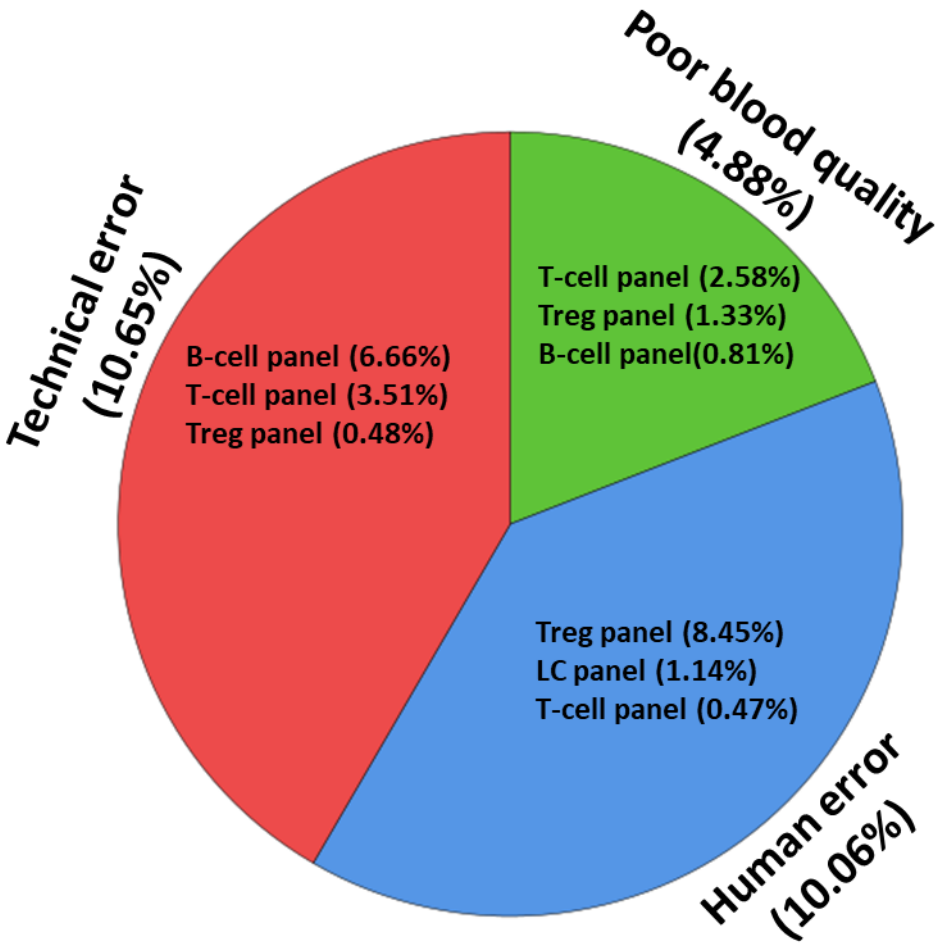


Figure 75 Errors across the four panels analysed.

Pie chart represent the 3 errors encountered during FC analysis. Red (Technical error), Blue (processing error), Green(poor blood quality)

Processing time of > 24 hours has been reported to impact data quality for all downstream assays (sometimes due to loss of viability) as well as flow cytometry analyses performed routinely [679, 741]. Lymphocyte enumeration was found to be overestimated in blood samples after >24 hours at room temperature before processing and staining [696, 741].

My findings suggest that all data retrieved from the NHS server were possibly allowing for one or more types of errors. This necessitated the development of exclusion rules that should be incorporated in the final SOP for the FC biomarker test. Combined (**figure 75)** processing error, faulty acquisition, poor sample quality, I identified a failure rate greater than 10% in data passing this NHS QA/QC procedure which was indeed much higher compared to the 2% failure rate for flow analysis of peripheral blood previously reported in another study [679]. This highlights that tests performed as research rather than NHS need to be considered carefully if used in clinical practice.

In more detail, for the 3 most important CD4+T subsets used in my overall PhD project (naïve, IRC, and Treg) gating remains the most subjective part of FC analysis. I re-analysed flow cytometry raw data from the T-cell and Treg panels while excluding all samples fulfilling exclusion rules. Results of the paired sample analysis between observers showed that the results for the naïve subset were relatively consistent with my observation. For IRC, this was also observed with greater variability notably when reporting low frequencies.

Regulatory T-cell (Treg), after the initial optimisation period (2013-15), were not relatively closely comparable with my results. The discrepancy at the later period may be also attributed to a lack of strict adherence to the use of the CD127 maker (as the 2nd most essential marker for Treg) due to the lack of availability of the

antibody for some time while not reporting this missing Abs as an issue. This highlights the need to also develop reporting rules and/or circulating information that a test (for example, Treg) cannot be performed when some reagents (e.g., the most essential makers, FOXP3 for Treg) are not available. Disagreement between observers in subsequent periods suggests that adherence to the gating protocol for Treg was not strictly followed by some staff involved in performing the assay during this period (not reporting shortage of reagent to their manager for example). The use of multiple NHS staff over this different period may have led to some of the variability in gating patterns detected.

The result of the analysis of the direction of changes suggests that data for the 3 subsets were relatively comparable (rho ≥ 0.6) between observers while much better for the naïve subsets showing the best correlation value (rho=0.9336) with a slope of the line best-fit matching to the line of complete identify. This suggests agreement in the gating performed by both observers despite a significant mean difference for naïve enumeration demonstrated using the BA plot (estimated at a -2%) suggesting small a negative bias in NHS analysis compared to me.

BA plot and test for bias confirmed that there was measurement bias between observers for IRC, however, this was mostly affecting low IRC frequencies. These are less biologically relevant and suggest that IRC results obtained by NHS may likely be usable by the clinician as it is comparable to those obtained by myself (as well as by my supervisor).

Treg showed a significant negative bias estimated at 0.4%. This suggests that most data reported by the NHS were underestimated compared to mine, which might have been explained in part by overestimated gating due to the absence of CD127 antibody although, excluding samples with no CD127 antibody, did not improve this

issue. Therefore a need for careful training of observers for this panel may be needed as its subjectivity is higher than for the other subsets.

The level of measurement bias for naïve (2%) and Treg (0.4%) between observers, suggests that these deviations may fall within the uncertainty of measurements, thus also suggesting that both gating patterns delivered acceptable results, provided exclusion of the faulty result is adhered to.

The current study, therefore, demonstrated the need to review the gating protocol for flow cytometry analysis in the NHS setting notably for IRC and Treg, and the need to adopt strict exclusion criteria to ensure the analysis and reporting of only acceptable FC data. This would ensure that accurate test result for patients is made available to improve clinical decisions while permitting a streamlined clinical laboratory operation, and harmonised results across multiple centres.

Therefore, I propose a few points that would need adding to NHS SOPs to support the use of FC biomarkers in clinical practice:

1. Specific handling guidelines for transportation for whole blood samples sent for immunophenotyping: The delay in sample delivery to the lab should be reconsidered and probably set at < 24 hours. Therefore, records of the time of blood taking to the delivery to the lab should be recorded. This could form the basis for further study to confirm the impact of transport time on individual blood cell immunophenotyping.

2. Optimisation of instrument performance with a lab protocol that could track reliability and reproducibility of setting (compensation). The protocol should also be revalidated in case of a change in markers (number and/or colour) and recompensated fully each time. This is usually in place but change in software version was overlooked.

3. The staining procedure performed when a key reagent (for example FoxP3 for Treg) is missing should be reported

4. A QA/QC of sensitive markers (i.e., CD62L) should define how acceptable/non-acceptable profiles “look” by asking an experienced 2nd opinion.

5.1.2 Flow cytometry validation using Dry tubes

In this study, I validated the feasibility of this novel DT technology to enumerate the 2 most important blood cell subset biomarkers (naïve and IRC CD4+T cells) used for the management of RA patients. I report the performance of 2 single DT, one containing a cocktail of 5- antibodies designed for quantifying naïve-CD4+T, and inflammatory-related cells-CD4+T (IRC-CD4+T cell) and the other, containing a 3-antibody cocktail designed for quantifying Treg using the Becton Dickinson dry coating technology. In validating reproducibility, DT results were correlated with the conventional wet tube (WT) data to compare the equivalence of the staining pattern and numerical data obtained using both technologies.

The DT data distribution was equivalent to WT for naïve but differed significantly for IRC. Although the profile was not absolutely identical, the staining pattern observed in DT for naïve and IRC allowed for a better subset separation, particularly for the IRC. The combination of fluorochromes in DT, notably for CD45RA showed a better separation with the BV421 fluorophore conjugate in DT than in the WT with PE conjugate. The DT panel design seems to have enabled a better separation of IRC than the WT design which could be advantageous in the NHS setting as discussed in the section, although the use of different flow cytometer machines for the two methods may have also played a role. DT were also designed as new with commercial constraint in the choice of clones and technical ones related to fluorochrome and the chemistry drying-technology, while the WT were designed over 15 years ago (2008 when more recent advance in FC functionalities) and maintained

for consistency over time/cohorts. The newly designed DT may therefore incorporate between reagents than those available in the past based on TRIMID groups' experiences and ensuring optimal detection of antigen expression, which may be fluorochrome dependent.

For the Treg panel, the staining was not detectable for the FoxP3- antibody (Alexa flour colour conjugate) in DT so far for an unknown reason. Data acquisition for a full Treg phenotype could not continue, although a "putative" Treg (defined by cell surface expression CD4/CD127/CD25) showed a comparable staining patterns. BD had reservation about the compatibility of intracellular staining and the chemistry of the DT, and while it was not fully validated in their lab, the putative Treg subset based on cell surface (CD127) was not an issue. More work is needed on both sides to resolve this issue and understand what is causing this lack of reactivity (i.e. whether FoxP3 specific or intra-cellular related).

DT versus WT comparison was satisfactory for naïve subset numeration with no evidence of bias between the two methods. This suggests that naïve quantification by the conventional WT routinely used in NHS is reproducible using the novel DT.
Although, IRC showed poor correlation and evidence of bias, the BA analysis showed that such observation was mainly affecting high frequency which was over-estimated using DT but not at low frequency. Low frequencies are considered within the healthy range and only high frequencies (above 2.5%) are relevant for RA management. This may therefore be a "good bias" allowing a better determination of IRC thanks to the change in fluorochrome for CD45RA clear increased frequencies will be easier to account for in a clinical decision.

Interestingly, the overall reliability test remained high for both naïve and IRC, despite better determination with DT for high frequency for IRC. My overall conclusion for this part of the work is that naive/IRC results obtained by WT will likely be reproducible if switching to DT technology.

DT workflow efficiency in the laboratory was also evaluated by comparing the turnaround time required to produce results.

Beyond technical improvement, DT allows for greater workflow efficiency and great applicability in terms of ease-to-use single-tube format [689, 690], **Figure 14** presents a streamlined workflow where time/risk for error of preparing antibodies cocktailing translated into reduced time to results thereby gaining about 30-40minute compared to WT. In the context of the screening of haematological malignancies, the use of the DT technology compared to liquid reagent also reported a significant gain of time and cost savings [688].

DT also presents a greater workflow standardisation as it eliminates the major sources of error arising from multiple manual pipetting. For instance, inadvertent omission of one or more antibodies as described before is one of the "processing" errors in NHS lab that can be difficult to discover (only visible at the later stage of acquisition) and may require a repeat analysis with increased cost [689] and further loss of time. The use of ready-to-use DT dry kit, therefore, could reduce day-to-day intra/inter staining variability between different samples and thereby lead to an increased result consistency across setting [690, 742].

Taken together, results show that the use of DT is equivalent to that of WT for the T-cell panel in terms of enumerating naïve and IRC subsets but not usable for a full Treg ($CD4^{+}CD127^{low}CD25^{high}FoxP3^{+}$) identification, while it may work for enumerating "putative" Treg which has been reported by other investigators [742]. The current study, therefore, demonstrated the feasibility of using the DT Technology

for T-cell panel as a replacement for WT flow cytometry method in the NHS routine immunology laboratory. The DT technology offers a powerful tool to simplify and standardize flow cytometry workflow and maximize laboratory efficiency by eliminating repetitive pipetting and human-prone errors, notably missing Ab due to handling mistakes. The DT will deliver easy-to-use standardization of multicolour flow cytometry data across multiple instruments and set the stage for its clinical application in routine NHS immunology service and in diverse areas of flow cytometry-based clinical trials/research worldwide.

5.1.3 Validation of FC technology using frozen samples

Immunophenotyping of immune cell subsets using frozen PMBC could allow for retrospective data collection and comparability of results in case of multicentre clinical settings. However, the use of frozen samples is associated with alteration in the integrity of the cell (viability), cell phenotype (influencing the frequency of cell enumeration), and/or cell functionality. The impact of freezing for a prolonged time on immunophenotyping in the context of biomarkers in inflammatory arthritis disease has been sparsely documented. This study, therefore evaluated the effect of freezing samples on the immunophenotyping of lymphocytes and their subsets.

Freezing PBMC (over 2-3 years storage) resulted in a substantial reduction in cell viability (10% loss of sample) with a possible risk of data loss while inducing a marker-specific alteration in the frequency of lymphocyte subsets (particularly CD62L) compared to fresh samples.

My data suggests that for lymphocytes and subsets enumeration, PBMC with greater than or equal to 15% cell viability still offer suitable quality allowing flow cytometry analysis with limited risk of mis-enumeration. This remains a very high rate of exclusion for sample analysis compared to a 2% exclusion rate reported for fresh

peripheral blood specimens in a review [679] and ~ 5% from my data related to blood quality.

PBMC as well as other cell types are known to be susceptible to the impact of the freezing process [743]. Fetal bovine serum plus 10% dimethylsulfoxide remains the best cryoprotectant in the freezing process for cells in immunology [744] and was used here by our Tissue bank. PBMC isolation and freezing in such standard cryoprotectant, loss of viability in my study is therefore over the expected. Several processes occur during freezing (i.e., intracellular ice crystal formation, the osmotic imbalance between the intracellular versus extracellular space, and the formation of channels by the residual unfrozen medium outside the cells) that damage the cells leading to this loss of cell viability [743-745].

Under these freezing conditions, the effect of freezing PBMC should be expected to be similar. TRIMID group before now have observed lower quality in FC FSC/SSC in samples from rheumatic patients compared to HC, although using fresh blood. It is therefore possible that different diseases may influence lymphocyte cryo-sensitivity in a unique fashion [746] notably in inflammatory diseases such as RA and PsA explaining the higher loss of sample I observed in this cohort.

Within the lineage panel, although there were no substantial differences in the percentage of CD4+T and CD8+T cells, B-cells, NK-cells, and NKT-cells were higher in frozen samples paired with fresh ones. This was a sample from same patients but split into two before freezing. Several studies have also reported the absence of a substantial effect on the frequency of CD4+T- and CD48+T-cells [724, 747-751], consistent with my data. While those studies used frozen samples for short storage periods (<12 months), the consistency with my data

using frozen samples over 2-3 years, suggests that the length of storage is less unlikely to impact T cells enumeration.

I observed a higher percentage of B-cells, NK-cells, and NKT-cells. B cells were higher in frozen PBMC in contrast to other studies reporting that distributions remained unaltered [750] or showed loss of B cells [752]. The result for NK cells has also remained contradictory across studies with some showing a trend towards lower frequency in frozen samples [748, 751, 753] while others found that NK cells were not impacted substantially [747, 750]. This suggests that lineage subsets are probably relatively stable although surface makers for lineage cells may also be sensitive to cryopreservation. Surface marker for NK cell identification (CD56) has been shown to be susceptible to freezing [723, 724]. In several disease conditions (such as HIV), standardization of analysis of frozen samples has been adopted [709, 741]. In addition, the use of different antibody clones may also interfere with such comparison.

For CD4+T-cell subsets (naïve/IRC/memory) enumeration, I observed a significantly lower frequency in naïve CD4+T cells paralleled with a higher frequency in IRC CD4+T cells in the frozen sample in paired matched fresh samples. The memory subset showed relatively higher frequencies however, it still showed good correlation results between fresh and frozen samples, suggesting that the impact of freezing on memory CD4+T was minimal compared to the substantial alteration observed in naïve/IRC subsets. I was able to attribute this effect to the loss of surface CD62L most likely through shedding [723, 724]. Similar results were observed for the CD8+T cell subsets at a lower level of impact than on CD4+T cells.

A study also showed significantly lower frequency for naïve CD8+T and CD4+T cells (defined with the same CD45RA+ and CD62L+ markers) in frozen compared to fresh samples [712] . Therefore, the alteration in the naïve/IRC subset associated with CD62L instability (due to protein shedding associated with cryopreservation) was also observed by others and is a limitation to using frozen samples when investigating such particular subset [712, 754]. The impact of freezing on the expression of several markers naïve (CD45RA+CD62L+) and "central memory" (CD45RO+CD62L+) T cells) used to categorize various subsets of both CD4+/CD8+ T-cells was also reported [712].

To overcome this issue, a 1-marker (CD45RA+) phenotype for enumerating naïve/memory T-cell subset in frozen samples was used to allow data acquisition. CD4+T naïve/memory cells showed no significant difference between frozen and fresh samples using this 1-marker gate. This suggests that loss/change of CD62L expression due to freezing leading to the negative skewness in CD4+T subsets enumeration using the conventional 2 markers (CD45RA/CD62L). This allowed me to overcome the issue, although at the expense of being able to also enumerate the IRC subset. Considering that IRCs are present in early RA and PsA but in a small proportion (except in a few cases with very high inflammation) and in the fresh sample I had data for, this may not be a major concern for this particular trial.

In contrast, the 1-marker approach in the CD8+T cells failed to fix the impact of freezing as the naïve subset was statistically significantly different than in fresh samples. This suggests that this approach would lead to over-estimation of naïve CD8+T cells in frozen samples. IRC frequencies are always much higher in CD8+T cells, so this information would be lost when using frozen samples.

Previous studies have reported that there was no significant alteration of B-subsets [755-758] whereas other studies have found substantial alteration in frozen PBMC compared with fresh [712, 759]. I observed that there were no significant changes in the frequency of memory B-cells, putative Breg, and plasmablasts in line with other studies [755-758] while the frequency of the naïve B subset was significantly lower in frozen samples compared with paired matched fresh samples. The reason for the effect on naïve B-cells was not determined and I could not propose a reason for this while not an isolated observation. This agreement/discrepancy with other studies may therefore be related to a diseases-specific impact rather than freezing alone.

For the Th17 panel, I did not have access to matched pairs (fresh + frozen). I used fresh sample for comparison with the frozen trial sample, both being PsA patients recruited from the same EAC cohort. I also compared the frequency of th17 cells from late PsA samples (fresh, n=24) for which I had data available from our group. All samples being associated with an inflammatory arthritis condition allowed me to account for the effect of PsA and inflammation or freezing cells. The results presented here also show significant changes in the frequency distribution of the surrogate CD4+Th17 phenotype in frozen samples. This suggests that the surface markers for Th17 identification are likely to be impacted by freezing. However, this remains to be confirmed using paired matched fresh samples.

Overall, the data presented show that the freezing of PBMC sample induce limited changes in the cell lineages which suggests that the use of frozen samples for the identification of these lymphocytes may be safe. However, the consistent alteration associated with their subsets (naïve/IRC) is a major concern in

inflammatory arthritis where definitive immunological features are investigated where these T-cell subsets have been found to provide biomarkers value across the disease continuum in RA particularly. Hence, the use of frozen PMBC may lead to underestimation of naïve CD4+T and overestimation of IRC CD4+T using the 2 marker gate. The CD45RA 1 marker gate identification can be used but at the cost of losing IRC data.

The data have implications for biomarker validation studies specifically, in the field of translational research in autoimmune inflammatory arthritis where large numbers of PBMC are cryopreserved. This process usually involves a long storage period and then a subsequent thawing process which may partly explain the loss in sensitive surface markers leading to lower comparability of results across multicentre research settings. Further studies may be necessary to develop novel tools for the assessment of blood cell biomarker stability in cryopreserved immune cells. Nonetheless, the use of fresh samples notably in multi-centre clinical trials has its challenges [714, 715]. Alternatively, it may be better to use flow specific freezing SMART tubes as previously reported by our group [5] to cryopreserve whole blood (hence requiring less work).

These are suitable for clinical trial as they allows cryostability of these sensitive subsets (including Treg) and notably allow for IRC with CD62L shedding to be quantified with no issue. However, there is a substantial cost for these tubes.

5.2 Clinical utility of FC biomarkers

5.2.1 Predicting outcome across the IAC

5.2.1.1 Alignment of my cohorts with previous studies at the clinical level

Demographic/clinical data from the various groups of patients I used were aligned with other cohorts at similar stages of progression in the IAC.

As expected, ACPA+, as a classification criterion for RA [302] with high specificity (~98%) and present in about 50-70% of RA [303, 304], but relatively absent in non-RA (up to 6% in some diseases) [305-307] was significantly higher in RA (67.8%) than non-RA (14.1%). Similarly, RF+ was higher in at-risk progressors and RA. Data underpinning the association between both autoantibodies and progression from at risk [189, 312, 760] and at the diagnostic stage of the IAC previously were reported [761-765]. So TRIMID groups' results were no different to similar cohorts.

EMS, TJC, SJC, and systemic inflammatory marker (CRP) showed significantly increased values in association with poor outcomes across IAC as previously reported [761, 766] . DAS28 (and its component element), were nonetheless significantly higher in RA but lower in MTX-induced remission. At the clinical level, each study cohort used in my PhD therefore is representative of the wider community of patients.

5.2.1.2 LS predicting progression across the IAC

It is important here to differentiate the predictive value from the dysregulation of a LC population in a disease. Indeed, many subsets are disturbed in the IAC (from at risk to remission) compared to health, but that is different from being predictive of an outcome. Furthermore, the modelling strategy chosen used a forward method which intend to select the best predictors amongst many variables, rather than highlighting changes in the data distribution of a variable. As such it shows that progression across the IAC is predicted by certain subsets and that this is stage specific while it does not diminish the valued of the observations that other subsets may be dysregulated in the disease (i.e. related to its pathophysiology) but not predictive. As such disturbance in Treg, IRC naive- CD4 T-cells, Breg, plasmablasts, NK cells, and CD8T-cells were observed in several stages but as detailed below, only predictive in specific ones and in combination with different clinical variable.

Altogether, my findings demonstrated the value of performing an extensive phenotyping of 18 LS to understand further the pathogenesis of RA, while prediction across the different IAC outcomes selected specific LS with a reduction in the number of LS needed to achieve prediction as the disease progressed while considering clinical parameters that were relatively normal at the pre-clinical stage where the model included mainly demographic and life-style predictors.

Progression to IA in ACPA+ at-risk individuals showed clear improvement by the addition of the newly identified LS over the 3 LS already known. The clustering analysis suggested that progressors were specifically associated with 8 LS profiles, cluster–III dominated by high CD4-IRC and cluster-II by high plasmablasts and Treg. Non-progressor (cluster-I) was characterised by high Treg and CD4-naïve paralleled with low frequencies of plasmablasts and CD4-

IRC. These profiles were not associated with any particular demographic and clinical data suggesting that they are independently regulated from any genetic, environmental or inflammatory events.

My findings, therefore, confirmed previous data for CD4+T-cell subsets and further identified 5 more LS associated with progression towards IA in ACPA+ at-risk individuals, providing clues to the identity of cells (CD8, NK, and B-cells) involved in the events triggering or associated with progression.

Importantly, different subsets were retained in the modelling for overall (**Treg, CD4-naïve, CD8, NK-cells, B-reg, and plasmablasts** and smoking, RF, HLA-SE, Tender-Joint-Count-78) compared to imminent **IRC, CD4-naïve, Treg, NK-cells, and CD8+T-cells** and smoking, RF) progression, suggesting a time frame for different biological events and/or triggers. Overall, data confirm that in ACPA+ at-risk individuals wider LS dysregulation precedes the development of clinical synovitis, while also providing increased accuracy over previous models with only CD4+T-cells subsets [5] .

Patients with clinical synovitis who later developed RA were characterised by significantly low naïve CD4+T and Treg with high CD4-IRC against those who progressed to other types of IA (non-RA), suggesting that dysregulation in CD4+T-cell subsets is not transient from at-risk phase but rather a persistent trigger for progression towards developing RA [5]. The possible value of these subsets to diagnose RA was established briefly by our group . Here, my extensive LS modelling as diagnostic biomarkers validated the predictive value of naïve CD4+T towards progression to ACPA+ RA from the IA stage while the other subsets have no added predictive value despite some of them being indeed dysregulated in RA (for instance, IRC CD4 cell, Treg CD4, and CD4 T-cells).

As in previous work [5], I also investigated the effect of autoantibodies (RF and ACPA) on the model because the EULAR 2010 criteria majorly depend on both autoantibodies for RA classification also considering the overarching OR of ACPA include in model 3 (OR 12.875, **Table 36**). In the current cohort (n=306), I constructed a combined Model (clinical without RF and ACPA + Flow data) this time using the forward approach. The model consistent with our previous work accurately predicted progression to RA in 75.16% of evolving IA patients while selecting only 3 variables: first, naïve CD4+T-cells with the highest contribution (30.8%), followed by SJC28 (26.0%), and finally age (11.9). This further strengthens the highly predictive value of naïve CD4+T in RA diagnosis regardless of the modelling approach used (enter or forward) particularly when not including APCA in the model as a classification marker.

A new classification biomarker for sero-negative RA is the most urgent need. With this consideration, I constructed a model (Model 4,**Table 36**) showing the specific LS profile associated with progression towards ACPA-negative RA, providing clues to the identity of cells (naïve CD4+T and CD4+T) involved in the events triggering progression towards ACPA-negative RA. The model predicted accurately 84% of ACPA-negative patients who developed RA with both high sensitivity (82.50%) and specificity (85.71%). However, external validation of this model will be needed using a larger replication sero-negative cohort.

Predicting MTX-induced remission at diagnosis will not only reduce cost in responders but also help gain time in averting the functional debilitating consequences of RA with early use of biologics in potential MTX non-responders. In patients with new onset RA, I tested the hypothesis that multiple dysregulations in circulating lymphocyte subsets homeostasis before 1st line treatment initiation

will predict with higher accuracy patients who are likely to respond to MTX while discriminating those who may never benefit from the use of the 1st line monotherapy. My data showed that MTX-induced remission was **only** predictive by high naïve CD4+T-cells. On the other hand, Treg, IRC and CD4 T-cells still showed slight difference in data distribution with no predictive added value over naïve CD4+T-cells and clinical data. Of note, modelling (fully adjusted for clinical parameters) used a forward method that allows the model to select the best predictive biomarkers confirmed high naïve CD4+T cell as the only LS retained for MTX-induced remission based on DAS28<2.6 with smoking and CRP retained in the model and 70.20% accuracy (AUC=0.734) .

Our group reported a significant value for a naïve CD4+T-cell subset being predictive for remission in ERA treated with MXT(n=120) although using the "Enter" method (other variables not contributing significantly to the model). I performed logistic regression using the "Forward" method to compare the previous model on my current data. Interestingly, both models retained the same set of predictors (smokers, CRP, and naïve CD4+T) regardless of the logistic regression method used with closely comparable accuracy. This confirms that a model with naïve CD4+T is indeed sufficient to predict MXT-remission while adjusted with demographic/clinical data. As a result, employing naïve CD4+T cells as a biomarker may have positive implications for a personalised treatment, as well as the cost-effectiveness of treatment in early RA patients starting MTX. However, this finding may need to be further validated by an independent study notably in early RA disease.

Prediction of response to drugs such as biologics may use different cell subsets as conventional synthetic disease-modifying anti-rheumatic drugs (csDMARD).

Both types of drugs potentially have different targets and may have different rates of response at different stages of the disease. In the first study by our group , higher naïve cell frequency was not found to be associated with remission in patients who received MTX concomitant to anti-TNF (n=70), suggesting a unique role of the naïve subset in the prediction of MTX response in RA. This remains to be examined in more details.

My findings showed an increase in circulating putative B-reg frequencies paralleling a reduction in Treg in pre-RA, suggesting an additional role for regulatory mechanisms before the onset of clinical synovitis. Functionally defective B-regs have been associated with homing to synovitis [270] while other studies have associated inflammation with the expansion of B-regs [767, 768]. In new-onset RA, the failure of B-reg to maintain a functionally suppressive Treg population was demonstrated [270, 769]. It is therefore conceivable that a decrease in Treg (loss of tolerance), paralleled with an increase in B-regs (subclinical inflammation) are both associated with progression to IA.

On the other hand, my findings do not exclude that, B-regs may be functionally defective [770, 771] or were excluded from the synovium (i.e., circulating), limiting their ability to perform their role locally. Future work will be needed to determine which hypothesis may be correct using functional studies to understand the reason for the increased Breg associated with at-risk progressors.

Consistent with our original work [3, 5], a reduction in circulating naïve and Treg CD4+T-cells frequencies and an increase in IRC-CD4 predates the development of IA . Since this original work, a standardised normalisation procedure for the naïve/memory and Treg subsets was established [3, 5], allowing analysis of different phase-specific outcomes across the IAC using continuous data.

Applying this to this group of 210 patients provided further validation of the original 3 CD4+T-cell subset model using the same logistic regression approach based on entering all variables in the model (enter method), however, it suggested that IRC-CD4 were no longer independently contributing to the prediction in this larger group (OR=1.104, p=0.144). Modelling using a forward method allowing for the best predictors only (accuracy=78.6%, AUC=0.880) confirmed that IRC was less predictive of the overall progression over 10 years, while still highly associated with rapid progression over 12 months, as also observed in the clustering **(Figure 47)** where most of the rapid progressors (68%) were in Cluster 3 defined by higher CD4+IRC.

We have associated many defects in the naïve CD4+T-cell subset with early RA pathogenesis, notably in relation to a decline in thymic activity, aberrant signalling, and aberrant proliferation reducing their TREC (T-cell receptor excision circle) content by more than 50% (1-2 cell cycle) [7], impaired IL7 responsiveness [772] and recently an IL6-driven network of epigenetic modifications suggesting the development of a subpopulation expressing more pro-inflammatory cytokines and closely resembling IRCs [1, 428]. In addition, naïve CD4+T-cell loss was also shown to directly result from the differentiation of naïve cells into IRC [7], a process driven by inflammation directly related to measures of inflammation [7] with a central role for IL6 in driving such changes and loss of IL6R expression as a result of its signalling [428]. Furthermore, IRCs persist when inflammation is subclinical due to reduced expression of pro-apoptotic genes (Bax expression) [1] and are associated with the occurrence of flares in patients in synthetic-DMARDs-induced remission [773]. Most importantly, IRCs remained naïve to an antigen challenge (hence expressing CD28) [1, 7] and were shown to be recent progeny

of naïve cells with a high content of T-cell receptor excision circles (TREC) [7]. As such, they are not to be confused with terminally differentiated T-effector memory T-cells (TEMRA) re-expressing CD45RA with controversies about high/low levels of expression of CD45RO and expression of CCR7 (reported negative or positive), while more consistently lacking CD62L, CD27 and CD28 expression, with also differences between CD4+ and CD8+ T-cells [774-777] and are antigen-experienced. The 5-colour panel used here does not include all the markers that could definitely differentiate all the various memory subsets of T-cells, but this does not alter the biomarker value of the IRC phenotype identified in this study. Indeed, here, we observed that IRC-CD4 enables segregation of a particular cluster of patients, the majority being rapid progressors, progressors being in Cluster-III. This further supports the hypothesis that imminent progression towards IA may be driven by an event involving or resulting from subclinical inflammation, driving or being associated with the differentiation of naïve CD4+T-cell into IRC as previously hypothesized [3, 7, 428].

Other CD4+T-cell subsets are relevant to RA pathology, notably Th17 cells and Tfh cells [778, 779]. However, it would have been very limiting to include panels for these in 2017-19 having started this work long before these were identified. Alternatively, the addition of a test for the enumeration of Th17 cells using a DNA-methylation specific qPCR [359], is however possible and should prove informative using stored whole blood. These subsets remain to be investigated more closely for a potential biomarkers value in the future.

Higher CD4+T-cell predated RA development in patients with clinical synovitis against those who developed other types of inflammatory arthritis (non-RA). In contrast, though later in established RA, a lower CD4+ T-cells have been reported in severe cases possibly (associated with intestinal bleeding) [780, 781].

Congruent with my result, another study comparing RA versus OA (a type of non-RA), reported significantly higher CD4+T-cells frequency in RA [782]. In contrast to no difference being reported in another one [783]. Therefore, the relative increase (~4%) in CD4+ T-cell in RA was visible probably because a variety of non-RA diseases (e.g., OA, SpA, pSA etc.) was included in the comparison group. Although CD8+T-cells share some of the pathways genetically associated with CD4+T-cells in RA [206], to my knowledge they have not been reported for their potential biomarker value, with one report presenting an association with arthralgia[464]. In preclinical IA, reports however suggest that CD8+T-cells make up approximately 40% of the total T-cells infiltrating the synovium [184, 559]. Following prolonged/chronic exposure to infectious agents, alterations in homing molecules expressed by CD8+T-cells occur [468], resulting in enhanced/altered migration. Here, I indeed demonstrated a reduced frequency of circulating CD8+T-cells, predictive of progression that could be reflecting migration/accumulation into the joint. Alternatively, this reduction could reflect a contraction of the CD8+T-cell pool following an infection (post activation cell-death), suggesting a role for subclinical infections as additional environmental risk for RA, as proposed in the past [473]. Furthermore, cell-death causes the release of inflammatory mediators that might also serve as triggers of cascades of events particularly Netosis [784]. In my study, NK-cells $CD56^{dim}$ were also predictive of progression suggesting that dysregulation of NK-cells may have a significant biological role, further fuelling an inflammatory cascade leading to disease progression. Several studies have indeed supported a role for NK-cells in RA pathology [293, 785] and at the onset of disease in ACPA+ RA and in pre-RA (arthralgia), although this observation was not directly related to progression [466].

Taken together, data presented in a flow chart (**see Figure 70**) showed that 3 novel cellular biomarkers (CD8, NK, and Breg) were identified. I demonstrated their predictive value for progression in addition to validating the predictive value of 3 CD4+T subsets across the IAC previously reported by our group.

- All 6 subsets were predictive for progression in seropositive + arthralgia individuals. (3 CD4+T cells were previously reported and currently validated by me while also showing the added value of 3 other subsets (CD8, NK, and Breg) not previously reported in this cohort)
- Only naïve was predictive for the diagnostic phase while Treg and IRC showed significant association. (previously reported by our group and currently validated in a large cohort in my thesis)
- For the clinical response phase, the naïve subset has shown predictive value for both csDMARDs- and bDMARDs-induced remission in RA-treated patients whereas only Treg and IRC were associated with remission in csDMARDs-treated patients alone (previously reported by our group and currently validated in a large cohort for csDMARDs only in my thesis).
- In flare, the 3 CD4+T subsets (naïve, Treg, and IRC) have shown predictive value but only with csDMARDs treatment while 2 of the CD4+T subsets (Treg and IRC) have been reported to be also associated with flare in patients treated with bDMARDs (previous report by the group).
- Of note, there were no identifiable biomarkers yet for seropositive individuals without arthralgia (hence bars with question marks) as it would require 10-year retrospective samples to be able to isolate such a cohort –a daunting task.

The major clinical benefit of being able to predict rapid progression would be for patients to access treatment at the critical very early IA/RA stage (i.e., 1-2 weeks of detectable synovitis) and to assess whether this can affect the long-term outcome/prognostic of these patients compared to routine early arthritis referrals (up to 2 years symptom duration). These are all studies currently ongoing in Leeds and hopefully will be reported in the future. Although few RA prevention trials have been reported (many being in progress/planned), it remains unclear whether IA development can be prevented in at-risk individuals and which drug/regimen would be most appropriate. Recent trials nonetheless suggest this may be possible: the APIPRRA study (ISRCTN-46017566 using abatacept [786] showed a sustained preventive effect at 2 years (data presented at EWRR and EULAR 2023); the PRAIRI study (using rituximab [12] also showed delayed progression. My data add valuable information that could contribute to risk stratification (as well as an understanding of the biology of the at-risk phase), but do not in themselves justify treatment for the time being. Different LS being indicative of the imminence of IA, our data may therefore find their best utility in selecting possible interventions targeting these cells/subsets/events, supporting personalised clinical-decision making, and guiding the selection of patients best suited for such preventive intervention.

On the practical side, although my group transferred these panels to the NHS services (back in 2013, based on the use of fresh whole blood samples), currently these specific panels are still only available in Leeds, while the technology itself is used worldwide and can provide data for clinical use in about 3-5 hours. A protocol that would allow for frozen blood samples (using SmartTubes^{TM}) [7]

could facilitate the use of these panels in a single flow centre (i.e., for retrospective analysis of samples as well as prospectively).

Alternatively, the technology can easily be replicated as recently shown in a collaborative work between Leeds and France (panels were transferred and data were replicated and merged with no bias) [9]. Careful planning of the number of antibodies/panels could be rationalised to suit local technical flow machine capacity (using SSC/CD4 gating for example). Novel technical development such as the use of DT technology whereby antibodies are pre-coated on flow plastic tube would also considerably help reduce procedure time (no pipetting) and increase adherence to SOP (although not suitable for Treg).

Using flow cytometry data as biomarkers in the NHS can be robust and feasible, despite technical issues, possible human error and subjectivity in gating strategies, and outweigh the challenges in view of the potential benefit to patients. The strength of FC as a biomarker tool lies in its ability to provide valuable information into disease prediction and management within a day (like most other laboratory test (including FC) generated by the NHS) that can guide personalized treatment strategies and improve patient outcomes. Furthermore, FC-based biomarkers are already largely utilised in NHS for many other purposes (see **section 1.9.3**) .These FC-biomarker panels were transferred to the NHS in Leeds in 2013 from being previously done in the Translational Research in Immune Mediated Inflammatory Diseases (TRIMID) group for over 5 years (since 2008), after demonstrating their clinical utility [5, 359, 520, 772, 781, 787], being replicated in another laboratory in France [9].

While others FC-based biomarker developed by other groups have also demonstrated value in the past as previously discussed in **section 1.7 .** These findings collectively emphasize the potential clinical relevance of FC-based cell

subset data as potential biomarkers in predicting disease outcomes and guiding therapeutic interventions. Nonetheless, it is important to acknowledge that these studies were not replicated (to date) or performed in routine hospital service, while ours repeatedly validated our initial results, is performed in routine settings and showed added values over clinical only data models. Furthermore, the strategy used for modelling allowed individual prediction (using an individual patient's probability to achieve a outcome) enabling personalised medicine.

Further advancements in FC technology and international standardization efforts have improved the reproducibility and reliability of FC data, enhancing its potential for clinical applications across sites. Initiatives such as the EuroFlow Consortium [690, 788] aim to standardize flow cytometry protocols and data analysis algorithms, thereby minimizing variability and facilitating the comparison of results across different laboratories and studies. Independently, the current study re-analysed the raw data retrieved from NHS and performed a rigorous validation analysis confirming the reliability of data produced by the NHS compared to data produced in research settings (ICC > 0.70). This suggests that with standardized protocols and rigorous training, the personal subjectivity of gating strategies can be mitigated to a large extent. Human error can also be mitigated by implementation of SOPs and the development of strict exclusion rules for poor samples quality.

Altogether, the rich information provided by cell subsets and the ongoing efforts to standardize protocols and improve analysis methods makes them worthy addition to the panel of tool with hight clinical utility for the management of RA. Thus, despite challenges, FC biomarker for cell subsets should become a adopted tool in personalised medicine and hold promise for improving patient care.

In conclusion, my study suggests that LS homeostasis is dysregulated at the early stage of the RA disease continuum before clinical synovitis occurs. I have demonstrated the additional predictive value of CD8+ T-cells B-reg and NK-cells besides the previously established CD4+T-cell subsets while further validating the highly predictive value of naïve CD4+T both in diagnosis and MTX-induced remission previously reported by our group. I demonstrated that perturbations in different subsets are associated with progression to arthritis, including rapid progression within 12 months, suggesting that additional time-dependent cell-based events are necessary for the progression to IA, while the development of systemic autoimmunity is not sufficient alone. Secondly, these panels are simple to perform routinely offering new tools to manage and stratify the risk of developing IA in ACPA+ at-risk individuals. However, there is still a need to investigate other LS not included in the 18 subsets reported here. For instance, Th17 and Tfh subsets that have shown to fuel inflammatory response in the synovium[211, 212, 223] and their hypothetical role in pre-clinical RA needs to be confirmed (which the pandemic prevented me to start). FC-based identification of cellular subsets is indeed a valid biomarker, but alone it may not capture the full complexity of diseases like RA and other inflammatory conditions. Other biomarker approaches such as genetics, transcriptomic analyses, proteomic profiling, and epigenetics may offer insights into the underlying mechanisms and can complement cellular subset analysis, enhancing our understanding and predictive models of these diseases. Integrating these various biomarker data can enhance the predictive power of models, aiding in diagnosis, prognosis, and treatment strategies.

5.2.2 Clinical utility of flow cytometry data using frozen samples

The original plan for this clinical trial (GOLMePsA) was to conduct the work at the NHS-immunology laboratory facilities in James University Hospital. However, due to COVID related restrictions to accessing laboratory space as well as the closure of clinics for over 18 months (all appointment were done by telephone), there was a need to change the original intended work plan. As a result, the flow cytometry work had to be carried out in the WTBB lab, utilizing frozen samples from only 31 patients in the clinical trial which luckily had PBMC stored at 2-time points baseline and 12 months. More samples were collected once clinic resumed but there was no time left to include them in my analysis (and no 12 months follow-up).

In the absence of drug arms and any other clinical information, I can only speculate about the meaning of these data. The trial was supposed to be randomized for treatment into equal halves, I used consecutive patients who had achieved the 12 months follow up. As such I was hoping to have equal representation of both arms (MTX and MTX+ golimumab). The difference between cluster A (n=22) and cluster B (n=7) being based on BL difference in B-cells subsets is unlikely to represent the drug arms as these would suggest a bias in randomisation. The absence of noticeable changes in subsets disturbance is more surprising and suggests very stable LS profiles in PSA which is not the case in RA. However, there is the possibility that it may associated with response and non-response to treatment as non-response in early PsA (~40%) is less frequent than in RA (~50%) [789-791]. Therefore, and highly speculatively, 22 versus 7 patients cluster groups could be related to response (n=22) versus non-response (n=7) as in early disease, non-response outcomes are not expected at a high frequency. At this point, I cannot speculate further until I have access to the

clinical data (sometime in summer 2024), unveiling of drug arms and responses group for me to validate this analysis. This will be part of further work, also adding data from 50 more patients who have now finished the study (December 2023).

While the original aim of investigating the LS for predicting treatment outcomes in early PsA patients could not be achieved due to COVID and MHRA/clinical trial unit restrictions on accessing clinical data (for unblinding/response). The LS data may still prove valuable for the overall trial analyses and the remaining FC data will be acquired. Over the work in the EAC on 18 LS for diagnostic purposes, 26 PsA patients were classified as non-RA. In these patients, the profile of LC changes was different from that seen in RA (as described below **Table 45**) as well as for Th17 from another study in fresh blood[9].

My data in frozen samples (at BL) are aligned with the disturbances specifically seen in PsA. For instance, NK cells have been reported to be reduced in number in PsA due to their recruitment to the synovium[792] .The progression from PsO to PsA is marked by disturbance in immune profiles, with specific T-cell subsets (CD8+T and Treg shift towards reduction) playing key roles in perpetuating inflammation[793, 794] . This highlights the pathological role of lymphocytes in during the course of the disease which may hold promise as biomarkers for predicting outcomes in PsA patients,

Table 45 LS analysis in 26 PsA patients compared to HC in fresh blood samples (BL data)

	HC (n=35)	PsA (n=26)	p value	*RA (n=206)
LS				
CD4%	48 (29; 66)	39.8 (17; 68)	0.008	55.29 (46.93-61.88
CD8%	17.9 (6; 43)	14.2 (6; 30)	ns	18.15 (12.72-23.39)
B %	10 (3.7; 19.3)	12.4 (4.5; 25.6)	ns	9.85 (7.62-13.64)
NK%	13.6 (5; 35)	8.7 (1.4; 35)	ns	7.61 (4.90-13.05
NKT%	3.4 (0.6; 27.4)	2.1 (0.1; 20)	ns	1.69 (0.87-4.10)
CD4+T-cell				
Norm-naïve $	-0.1 (-19.7; +142)	+2.2 (-22.8; +17.4)	ns	-4.71 (-13.81-4.75)
Norm-memory $	+0.6 (-10.6; +16.1)	+0.5 (-7.4; +18.3)	ns	-3.51 (-7.59-1.34)
IRC	1.3 (0.1-7.9)	2.8 (0.6;29)	<0.0001	2.80 (1.30-6.00)
CD8+T-cell				
Norm-Naïve $	-3.3 (-23; +22.5)	+9.4 (-18.8; +39)	ns	5.76 (-4.07-13.91)
Norm-memory $	+0.1 (-21.3; +28)	-16.6 (-23; +44)	<0.0001	-16.85 (-25.45- -7.77)
IRC	1.0 (0.5 ; 36)	7 (2 ; 28.5)	<0.0001	16.25 (7.45-23.70)
B-cell				
Norm-naïve $	-3.5 (-28; +20.2)	-6.1 (-37; +13.3)	0.042	-0.99 (-13.92-10.11)
Norm-memory $	-5 3 (-23; +31)	+2.9 (-22.6; +41)	ns	-2.595 (-11.29-13.24)
Breg	2.5 (0.5; 5.8)	+4.7 (0.4; 13.6)	<0.0001	4.10 (1.60-5.70)
Treg				
Norm-Treg % $	+0.1 (-2.8; +2.8)	-0.6 (-2; +1.3)	ns	-1.64 (-2.88- -0.50)
Th17 cells				
Th17 %	6.5 (4.1; 13)	5.1 (1.2; 14)	ns	Not done

$ normalised values, , data are presented as median (range. *RA result from EAC cohort (**see Table 34, page 274**), ns=non-significance

Conclusion

The overall hypothesis of my thesis was that lymphocyte dysregulation assessed by flow cytometry could provide biomarkers for the management of rheumatoid arthritis (RA). The first part of this hypothesis was achieved.

1. assessing the reliability of the wet-based FC data obtained by NHS services,
2. validating the use of novel flow-cytometer dry-tube technology compared to wet-tube currently used, and
3. finally validating FC on frozen samples

In connection with the first hypothesis, I showed that wet-based FC data retrieved from the NHS were affected by one or more types of errors. When combined (Human error-missing Ab, Technical error-faulty acquisition, Poor sample quality), resulted in a failure rate greater than 10% for generating data passing the QA/QC procedure required for FC biomarker test. This necessitated the development of rules for exclusion that should be incorporated in the final SOP for a flow cytometry biomarker test. The study further demonstrated the feasibility of using the novel DT Technology for T-cell panel as a replacement for the WT flow cytometry method in the NHS routine immunology laboratory. For the FC validation on frozen samples, I demonstrated that only PBMC with ≥15% cell viability represented a suitable viability limit that allowed flow cytometry analysis with limited risk of mis-enumeration.

As a second aspect, the involvement of immune cells in RA pathogenesis has been extensively studied [5, 7, 428, 583, 726-729], leading to a compelling need to identify comprehensively blood cell flow cytometry biomarkers with predictive value that can support improved RA management; from at-risk individuals, to

diagnosis of RA, and selecting optimal treatment option with improved accuracy. My group has demonstrated the predictive value of a few cellular biomarkers (notably multiple CD4+T subsets) for progression across the IAC, however, the effector profile of other immune cell types in all stages of the disease across IAC is yet to be fully explored. To bridge this gap, the second part of my thesis set out to test the hypothesis.

1. That comprehensive blood cell phenotyping has added clinical value for the management of RA:
    (i) in predicting clinical outcomes across the IAC, and
    (ii) ii) providing further understanding of the events driving progression.

This aspect was also achieved. My data demonstrated that in ACPA+ at-risk individuals wider LS dysregulation preceded the development of clinical synovitis, while also providing increased accuracy over previous models with only CD4+T-cells subsets. I also demonstrated the additional predictive value of CD8+ T-cells, B-reg, and NK-cells besides the previously established CD4+T-cell subsets while further validating the highly predictive value of naïve CD4+T both in diagnosis and MTX-induced remission previously reported by our group. Therefore, my PhD provided further insight into understanding cellular events driving progression across the IAC and offered clinicians a set of new tools to manage and stratify the risk of developing Inflammatory Arthritis in at-risk individuals.

Finally, I demonstrated a new set of classification biomarkers for sero-negative RA, providing clues to the identity of cells (notably, naïve CD4+T and CD4+T) involved in the events triggering progression towards ACPA-negative RA patients with a predictive accuracy of 84%.

In addition, DT technology offers a powerful tool that will deliver easy-to-use standardization of multicolour flow cytometry data across multiple instruments

and maximize laboratory efficiency by eliminating repetitive pipetting and human-prone errors.

5.3 Future work

I am hoping to complete the statistical modelling of outcome analysis and prediction for response to drugs in PsA patients from the GOLMePsa clinical trial which I could not finish due to the COVID-19-associated delay that affected the unblinded date. I will start as soon as the unblinding of the trial is done and clinical data is made available (sometime in 2024) to emerge with flow cytometry data (LS which is now fully analysed). Secondly, I will, together with my supervisor (Frederique Ponchel) and other clinical colleagues (Gabriele De Marco, another PhD student working on the trial, and Helena Marzo-Ortega, PI), write up the findings for publication.

To test the hypothesis generated from my thesis regarding Breg functionality; the increased Breg associated with at-risk progressors may be functionally defective or due to exclusion from the synovium, limiting their ability to perform their role locally. Future work will be needed to determine which hypothesis may be correct. Finally, to understand the role of NK/CD8+T in pre-RA, I would like to add other different subsets, Th17 and Tfh in pre-RA.

The study on the EAC cohort investigating the LS as biomarkers for the predicting RA diagnosis used a disease comparison group (termed the non-RA group) which comprised all other forms of inflammatory arthritis (PsA, ReA, OA, Gout, UA and AS as well as a few connective tissue disease) and non-persistent symptoms, In the future, it will be interesting to consider disease comparison groups independently as it would partly shed more light on the ongoing debate on whether LS may also have value in these disease. The TRIMID research

group already described such potential in AS for response to anti-IL17 but not for anti-TNF, there are clear disturbances in PsA (described above) as well as those observed in OA [4]. The value of these disturbances remains to be investigated to establish if any may prove to also have biomarker and novel programmes should be developed for this disease as did for RA.

Clinical predictors in the at-risk models developed in my PhD were slightly at variance with those from a previously published model [312] probably because different patients and parameters were used (notably 78-TJC rather than hands small TJC and importantly, ACPA-positivity validated by a research CCP2 test). Model-1 notably only explains 25% of the variance in predicting IA. Modelling using combined data in model-3 showed that 6 LS and the same 4 clinical parameters had superior value over model-2, and clearly over model-1 (+15.7% accuracy). Nonetheless, only 62% of the variance can be explained by model-3, leaving room for additional biomarkers to be added, possibly using imaging [795], genetics, transcriptomics, and cytokines previously discussed **section 1.6** or epigenetic modification as recently evidenced in early RA [428].

Enumeration of cellular subsets is indeed a valid biomarker, but alone it may not capture the full complexity of diseases like RA. Other biomarker approaches such as genetics, transcriptomic analyses, proteomic profiling, and epigenetics may offer insights into the underlying mechanisms and can complement cellular subset analysis, enhancing our understanding and predictive models of these diseases. However, to achieve a more complete view of the events that underpin progression and response to treatment in RA, it is likely that biomarkers present in several data types will need to be measured in the same individuals permitting a more holistic analysis approach. These various biomarker modalities complement each other and provide a more comprehensive understanding of disease mechanisms and heterogeneity. Integrating these various biomarker data can enhance the predictive power of models, aiding in diagnosis, prognosis, and targeted therapies tailored to individual patients' needs.

Furthermore, the presence of comorbidities in rheumatoid arthritis (RA) as previously mentioned in section Error! Reference source not found. can s ignificantly impact the identification and enumeration of cell subsets, although in the work described in this thesis, this is a relatively minor concern as the disease is in its pre-clinical/early stage and those have not yet developed. Cardiovascular disease (CVD), is often comorbid with long lasting RA. It develops due to the burden of chronic inflammation and may potentially alter the proportions of different immune cells. Specifically, CVD can induce alterations in immune cell subsets, particularly atherosclerosis-related changes. For example, plaque formation in blood vessels was associated with alterations in monocyte and T cell subsets, affecting immune cell distribution and function [796, 797]. Additionally, conditions like obesity (reflected in BMI) are linked to chronic low-grade inflammation, which can disrupt immune cell distribution-phenotypic and function changes through various mechanisms like metabolic dysregulation [798]. Diabetes Mellitus (DM) is another common comorbidity in RA, and both diseases share inflammatory pathways [799]. Here hyperglycemia and insulin resistance can dysregulate immune cell function, affecting subsets T cells and macrophages. Moreover, DM-associated complications like diabetic neuropathy and nephropathy can exacerbate inflammatory responses [799].

Understanding these late disease comorbidities is crucial for RA management, as they can influence disease severity and treatment outcomes [509]. Addressing these interconnected health issues holistically improves patient care and enhances overall prognosis in individuals with RA.

Further comparison with other models [350, 581] notably the most recent one [795] that used more than 30 candidates is difficult as they did not use the same statistical approach (also not providing AUC) but my model 3 showed high performance (AUC=0.911) for the overall prediction, while rapid progression is slightly less good but still high (AUC=0.794). On the other hand, I observed similar findings in RF+ only patients (4/13 progressors) which also confer risk for IA development in patients with arthralgia [6]. This was verified for naïve, Treg, IRC, and CD8 but not for NK and the B-cell subsets (possibly due to small numbers). This nonetheless suggests potential predictive value for these LS across at least 2 risk-related autoantibodies. Further work would be needed to evaluate these with respect to other risk factors such as genetic (HLA-DR SE) and/or lifestyle (smoking), however, I already showed higher predictive value for LS than for known risk factors.

Other limitations of my work include the relatively low number of subjects with all flow panels across the 3 stages of the IAC, notably due to B- and LC-panels not being available and the selection of ACPA+ participants with a highly specific CCP-2 test for the at-risk stage.

I also acknowledge the limitation of the statistical modelling approaches and performed the optimism correction to account for this in model-3 and 6 for the at-risk cohort, while the model with only the 3 CD4+T-cell subset clearly reproduced previous data [3, 5]. External cohorts would be critical to fully validate these findings. The main hurdle in any biomarker research programme is that it can only be replicated if the selection criteria for the study population are the same. For instance, there are many cohorts of at-risk individuals worldwide [312, 352, 765, 800, 801] but they all use different criteria to define individuals at-risk of RA,

preventing the generalisation of any findings. The data presented here nonetheless replicate previous models, while the final model using more subsets showed clear improvement across the IAC.

Addressing missing cell data was crucial for obtaining reliable estimates in the analysis. To handle missing data effectively, multiple imputation was employed, a method known for reducing imprecision and bias compared to other approaches like complete case analysis or single-value imputation methods [802, 803]. Unlike single-value imputation methods, multiple imputation does not replace missing data with a single predicted value. Instead, it inputs multiple values for each missing item, incorporating random errors to better reflect the true distribution of the original measure. By imputing missing entries based on statistical characteristics of the dataset, such as associations among variables and their distributions, multiple imputation provides accurate estimates and reduces the risk of false conclusions [803, 804].

Data cleaning was carried out to mitigate bias due to missing data and ensure a usable dataset was included in the statistical analysis and modeling. For instance, samples missing all B-cell and lineage subsets were excluded from the analysis before imputation (described in detail in section **4.2.1.1)**. This allowed multiple imputation to be performed for data with no more than a certain threshold of missing values, avoiding potential bias and loss of statistical power. Multiple imputation is a widely accepted statistical method for handling missing data, applicable to various types of data [805-807] , including cell data [795, 808-810]. However, its validity relies on assumptions about the mechanism of missingness, particularly that data are missing at random (MAR) [804, 811] . Unfortunately, the missingness mechanism is not usually fully known and is often a combination of more than one mechanism. However, including all possibly relevant predictors in

the imputation model [812-816] and conducting tests to possibly reject other mechanisms for missingness, such as missing completely at random (MCAR), as was performed in this work, ensured that the imputation model yield reliable results. Diagnostic checks also confirmed the validity of the imputed dataset. No significant differences obtained when the original dataset was compared to the imputed dataset suggest that the imputed data accurately reflected the original dataset.

Unfortunately, COVID-19 prevented the development of the Th17 panel within the NHS immunology service, hence further work will need to add Th17 cells to provide a more comprehensive picture and maybe as well as Tfh subset and overall refined the B-cell subset for certain class of drugs.

6.1 Covid Statement

The COVID-19 lockdown and the restriction policies affecting access to facilities both within the NHS and University settings were a major setback for my Ph.D. This had a major impact on my activities when working with human samples and data analysis as well as for using the offline flow cytometry facilities.

To summarise this, the covid-19 lockdown in March 2020, started just 5 months after my start date on 1st Oct 2019 and lasted for 6 months. Access to the lab remained limited for another 9 months after that.

My supervisor managed to provide me with alternative means to perform my data analysis for raw data already retrieved from NHS services before the lockdown (ie., FCS files from servers). However, this took 4 months and impacted my progress in acquiring data due to having to identify a suitable laptop with Windows 7 manager right to restriction to install DIVA software. This time was thankfully used to progress my literature reading and writing of my transfer report, although with very limited data could be included. I have unfortunately not been able to recover the lost time on this part of my PhD project.

Another aspect of my experimental work requires me to access NHS immunology premises which were also halted in March 2020 and to patients samples. This restriction has not been lifted until clinics resumed in late 2022. This also greatly impacted on the DT chapter of my thesis.

A 3rd aspect of my project included work on a clinical trial that was stopped due to MHRA restriction during the pandemic. This trial could not terminate as planned (end of 2021) in time for the drug unbinding and clinical data release that would be necessary for me to complete my analysis for this chapter of my thesis. This was the biggest challenge as access to clinical data (for a small set of 32 patients) was refused by the MHRA. With the support of my supervisors, we had

to design an alternative strategy for me to exploit the flow data acquired and this took out a great part of my thesis as well. Thankfully, this time was used to develop additional data analytical and statistical skills.

References

1. Burgoyne, C.H., et al., *Abnormal T cell differentiation persists in patients with rheumatoid arthritis in clinical remission and predicts relapse.* (1468-2060 (Electronic)).
2. Gul, H., F. Ponchel, and P. Emery, *OP0182 IN RA PATIENTS IN REMISSION, WHICH BIOMARKERS PREDICT SUCCESSFUL TAPERING OF CSDMARDS?* 2021, BMJ Publishing Group Ltd.
3. Hunt, L., et al., *T cell subsets: an immunological biomarker to predict progression to clinical arthritis in ACPA-positive individuals.* Annals of the Rheumatic Diseases, 2016. **75**: p. 1884-1889.
4. Ponchel, F., et al., *Changes in peripheral blood immune cell composition in osteoarthritis.* Osteoarthritis and cartilage, 2015. **23**(11): p. 1870-1878.
5. Ponchel, F., et al., *T-cell subset abnormalities predict progression along the Inflammatory Arthritis disease continuum: implications for management.* Scientific reports, 2020. **10**(1): p. 3669-3669.
6. Ponchel, F., et al., *Added value of multiple autoantibody testing for predicting progression to inflammatory arthritis in at-risk individuals.* RMD open, 2022. **8**(2): p. e002512.
7. Ponchel, F., et al., *Dysregulated lymphocyte proliferation and differentiation in patients with rheumatoid arthritis.* Blood, The Journal of the American Society of Hematology, 2002. **100**(13): p. 4550-4556.
8. Ponchel, F., et al., *CD4+ T-cell subsets in rheumatoid arthritis.* International Journal of Clinical Rheumatology, 2012. **7**(1): p. 37.
9. Rabin, T., et al., *Association of circulating lymphocyte subsets with response to IL17i and TNFi in axial spondyloarthritis.* Clinical and experimental rheumatology, 2022.
10. Bugatti, S., et al., *B cells in rheumatoid arthritis: from pathogenic players to disease biomarkers.* Biomed Res Int, 2014. **2014**: p. 681678.
11. Robinson, W.H. and R. Mao, *Biomarkers to guide clinical therapeutics in rheumatology?* Curr Opin Rheumatol, 2016. **28**(2): p. 168-75.
12. Gerlag, D.M., et al., *Effects of B-cell directed therapy on the preclinical stage of rheumatoid arthritis: the PRAIRI study.* Annals of the rheumatic diseases, 2019. **78**(2): p. 179-185.
13. Chandrashekara, S., *Current studies of biomarkers for the early diagnosis of rheumatoid arthritis.* Current Biomarker Findings, 2014: p. 107-119.
14. Guo, Q., et al., *Rheumatoid arthritis: pathological mechanisms and modern pharmacologic therapies.* Bone research, 2018. **6**(1): p. 15.
15. Anaya, J.-M., et al., *Autoimmunity: from bench to bedside [Internet].* 2013.
16. Martin, C.R. and V.R. Preedy, *Scientific basis of healthcare: arthritis*. 2012: CRC Press.
17. Khare, S., *Unusual Oral Manifestations of autoimmune disorders.* Journal of Orofacial Research, 2023: p. 13-21.
18. Almoallim, H., et al., *A review of the prevalence and unmet needs in the management of rheumatoid arthritis in Africa and the Middle East.* Rheumatology and therapy, 2021. **8**: p. 1-16.
19. Linos, A., et al., *The epidemiology of rheumatoid arthritis in Rochester Minnesota: a study of incidence, prevalence, and mortality.* American journal of epidemiology, 1980. **111**(1): p. 87-98.

20. Symmons, D., et al., *The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century.* Rheumatology, 2002. **41**(7): p. 793-800.
21. Abhishek, A., et al., *Rheumatoid arthritis and excess mortality: down but not out. A primary care cohort study using data from Clinical Practice Research Datalink.* Rheumatology, 2018. **57**(6): p. 977-981.
22. Galloway, J., et al., *Direct healthcare resource utilisation, health-related quality of life, and work productivity in patients with moderate rheumatoid arthritis: an observational study.* BMC Musculoskeletal Disorders, 2021. **22**: p. 1-11.
23. UK, A.R., *State of musculoskeletal health 2017*. 2017, Arthritis Research UK London.
24. NRAS. *What is RA*. 2022 2 December 2023]; Available from: https://nras.org.uk/resource/what-is-ra/.
25. NICE. *Rheumatoid arthritis in adults: management*. 2020 10 November 2023]; Available from: https://www.nice.org.uk/guidance/ng100/chapter/Context#:~:text=The%20incidence%20of%20the%20condition,ages%20can%20develop%20the%20disease.
26. Yelin, E., et al., *The impact of functional status and change in functional status on mortality over 18 years among persons with rheumatoid arthritis.* The Journal of Rheumatology, 2002. **29**(9): p. 1851-1857.
27. van der Helm-van, A.H. and A. Zink, *What is rheumatoid arthritis? Considering consequences of changed classification criteria*. 2017, BMJ Publishing Group Ltd. p. 315-317.
28. Hazes, J.M. and J.J. Luime, *The epidemiology of early inflammatory arthritis.* Nature Reviews Rheumatology, 2011. **7**(7): p. 381-390.
29. Van Aken, J., et al., *Comparison of long term outcome of patients with rheumatoid arthritis presenting with undifferentiated arthritis or with rheumatoid arthritis: an observational cohort study.* Annals of the rheumatic diseases, 2006. **65**(1): p. 20-25.
30. Machado, P., et al., *Multinational evidence-based recommendations on how to investigate and follow-up undifferentiated peripheral inflammatory arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative.* Annals of the rheumatic diseases, 2011. **70**(1): p. 15-24.
31. Prete, M., et al., *Extra-articular manifestations of rheumatoid arthritis: an update.* Autoimmunity reviews, 2011. **11**(2): p. 123-131.
32. Allen, A., S. Carville, and F. McKenna, *Diagnosis and management of rheumatoid arthritis in adults: summary of updated NICE guidance.* Bmj, 2018. **362**.
33. Yap, H.-Y., et al., *Pathogenic role of immune cells in rheumatoid arthritis: implications in clinical treatment and biomarker development.* Cells, 2018. **7**(10): p. 161.
34. Black, R.J., et al., *Global, regional, and national burden of rheumatoid arthritis, 1990–2020, and projections to 2050: a systematic analysis of the global burden of disease study 2021.* The Lancet Rheumatology, 2023. **5**(10): p. e594-e610.
35. Black, R.J., et al., *Mortality estimates and excess mortality in rheumatoid arthritis.* Rheumatology, 2023: p. kead106.

36. Johnson, T.M., et al., *A Narrowing Mortality Gap: Temporal Trends of Cause‐Specific Mortality in a National Matched Cohort Study in US Veterans With Rheumatoid Arthritis.* Arthritis care & research, 2023. **75**(8): p. 1648-1658.
37. Vegas, L.P., et al., *Prevalence, mortality, and treatment of patients with rheumatoid arthritis: A cohort study of the French National Health Data System, 2010–2019.* Joint Bone Spine, 2023. **90**(1): p. 105460.
38. Almutairi, K.B., et al., *Mortality Trends Among Patients with Rheumatoid Arthritis in Western Australia.* Rheumatology and therapy, 2023: p. 1-17.
39. Yu, F., et al., *Secular trend of mortality and incidence of rheumatoid arthritis in global, 1990–2019: an age period cohort analysis and joinpoint analysis.* BMC Pulmonary Medicine, 2023. **23**(1): p. 356.
40. Soussi, B.G., et al., *Temporal trends in mortality in patients with rheumatoid arthritis: a Danish population-based matched cohort study.* Rheumatology, 2023: p. kead325.
41. Fazal, S.A., et al., *A clinical update and global economic burden of rheumatoid arthritis.* Endocrine, Metabolic & Immune Disorders-Drug Targets (Formerly Current Drug Targets-Immune, Endocrine & Metabolic Disorders), 2018. **18**(2): p. 98-109.
42. Sokka, T., et al., *Work disability remains a major problem in rheumatoid arthritis in the 2000s: data from 32 countries in the QUEST-RA study.* Arthritis research & therapy, 2010. **12**: p. 1-10.
43. Katchamart, W., et al., *Health-related quality of life in patients with rheumatoid arthritis.* BMC rheumatology, 2019. **3**(1): p. 1-8.
44. Malm, K., et al., *Quality of life in patients with established rheumatoid arthritis: a phenomenographic study.* SAGE open medicine, 2017. **5**: p. 2050312117713647.
45. Bąk, E., et al., *Factors that affect the assessment of the quality of life of rheumatoid arthritis patients depending on the prevalence of frailty syndrome.* Health and quality of life outcomes, 2020. **18**: p. 1-9.
46. Matcham, F., et al. *The impact of rheumatoid arthritis on quality-of-life assessed using the SF-36: a systematic review and meta-analysis*. in *Seminars in arthritis and rheumatism*. 2014. Elsevier.
47. Bai, B., et al., *Quality of life and influencing factors of patients with rheumatoid arthritis in Northeast China.* Health and quality of life outcomes, 2020. **18**(1): p. 1-10.
48. Ramírez, L., C. Rodríguez, and M. Cardiel, *Burden of illness of rheumatoid arthritis in Latin America: a regional perspective.* Clinical rheumatology, 2015. **34**: p. 9-15.
49. Newhall-Perry, K., et al., *Direct and indirect costs associated with the onset of seropositive rheumatoid arthritis. Western Consortium of Practicing Rheumatologists.* The Journal of rheumatology, 2000. **27**(5): p. 1156-1163.
50. Michaud, K., et al., *Direct medical costs and their predictors in patients with rheumatoid arthritis: a three‐year study of 7,527 patients.* Arthritis & Rheumatism, 2003. **48**(10): p. 2750-2762.
51. Van Hoogmoed, D., et al., *Physical and psychosocial correlates of severe fatigue in rheumatoid arthritis.* Rheumatology, 2010. **49**(7): p. 1294-1302.
52. Gwinnutt, J.M., et al., *Predictors of presenteeism, absenteeism and job loss in patients commencing methotrexate or biologic therapy for rheumatoid arthritis.* Rheumatology, 2020. **59**(10): p. 2908-2919.

53. NRAS. *Invisible disease: Rheumatoid arthritis and chronic fatigue survey* 2014 7 November 2023]; Available from: https://www.nras.org.uk/invisible-disease-rheumatoid-arthritis-and-chronic-fatigue-report.
54. Burton, W., et al., *Systematic review of studies of productivity loss due to rheumatoid arthritis.* Occupational Medicine, 2006. **56**(1): p. 18-27.
55. Birnbaum, H., et al., *Societal cost of rheumatoid arthritis patients in the US.* Current medical research and opinion, 2010. **26**(1): p. 77-90.
56. Sullivan, P.W., et al., *Influence of rheumatoid arthritis on employment, function, and productivity in a nationally representative sample in the United States.* The Journal of rheumatology, 2010. **37**(3): p. 544-549.
57. Houssien, A., et al., *SAT0098 The association between work disability and mental health in rheumatoid arthritis*. 2017, BMJ Publishing Group Ltd.
58. Peterson, S., et al., *SAT0089 Anxiety and Depression among Rheumatoid Arthritis Patients, and Association with Clinical Measurements and Patient Reported Disability and Treatment Satisfaction*. 2016, BMJ Publishing Group Ltd.
59. Minnock, P., et al., *Factors that influence fatigue status in patients with severe rheumatoid arthritis (RA) and good disease outcome following 6 months of TNF inhibitor therapy: a comparative analysis.* Clinical Rheumatology, 2015. **34**(11): p. 1857-1865.
60. Gulacsi, L., et al., *Biosimilars for the management of rheumatoid arthritis: economic considerations.* Expert review of clinical immunology, 2015. **11**(sup1): p. 43-52.
61. Emery, P., et al., *A European chart review study on early rheumatoid arthritis treatment patterns, clinical outcomes, and healthcare utilization.* Rheumatology international, 2015. **35**: p. 1837-1849.
62. Nikiphorou, E., et al., *Association between rheumatoid arthritis disease activity, progression of functional limitation and long-term risk of orthopaedic surgery: combined analysis of two prospective cohorts supports EULAR treat to target DAS thresholds.* Annals of the rheumatic diseases, 2016. **75**(12): p. 2080-2086.
63. Taylor, P.C., et al., *Targeted literature review of current treatments and unmet need in moderate rheumatoid arthritis in the United Kingdom.* Rheumatology, 2021. **60**(11): p. 4972-4981.
64. Kitas, G.D. and S.E. Gabriel, *Cardiovascular disease in rheumatoid arthritis: state of the art and future perspectives.* Annals of the rheumatic diseases, 2011. **70**(1): p. 8-14.
65. Hughes, C. and N. Gullick, *AB0303 Does comorbidity adversely impact on treatment response in patients with rheumatoid arthritis*. 2017, BMJ Publishing Group Ltd.
66. Solomon, D.H., et al., *Risk of diabetes among patients with rheumatoid arthritis, psoriatic arthritis and psoriasis.* Annals of the rheumatic diseases, 2010. **69**(12): p. 2114-2117.
67. Liao, K.P. and D.H. Solomon, *Traditional cardiovascular risk factors, inflammation and cardiovascular risk in rheumatoid arthritis.* Rheumatology, 2013. **52**(1): p. 45-52.
68. Leggett, S.A., et al., *159. THE IMPORTANCE OF ACHIEVING CLINICAL RESPONSE TO TREATMENT AND CHANGES IN PHYSICAL ABILITY AND QUALITY OF LIFE ON WORKER PRODUCTIVITY OUTCOMES IN RHEUMATOID ARTHRITIS: RESULTS FROM THE BRITISH SOCIETY*

FOR RHEUMATOLOGY BIOLOGICS REGISTER FOR RHEUMATOID ARTHRITIS. Rheumatology, 2017. **56**(suppl_2): p. kex062. 160.
69. Galloway, J., et al., *The impact of disease severity and duration on cost, early retirement and ability to work in rheumatoid arthritis in Europe: an economic modelling study.* Rheumatology advances in practice, 2020. **4**(2): p. rkaa041.
70. Betts, K.A., et al., *Economic burden and treatment patterns of cycling between conventional synthetic disease-modifying antirheumatic drugs among biologic-treated patients with rheumatoid arthritis.* Clinical therapeutics, 2016. **38**(5): p. 1205-1216.
71. Zhang, W. and A.H. Anis, *The economic burden of rheumatoid arthritis: beyond health care costs.* Clinical rheumatology, 2011. **30**: p. 25-32.
72. Control, C.f.D. and Prevention, *National and state medical expenditures and lost earnings attributable to arthritis and other rheumatic conditions--United States, 2003.* MMWR. Morbidity and mortality weekly report, 2007. **56**(1): p. 4-7.
73. Badley, E.M., *Arthritis in Canada: what do we know and what should we know?* The Journal of Rheumatology Supplement, 2005. **72**: p. 39-41.
74. Maravic, M., *Economic impact of rheumatoid arthritis (RA) biotherapies in France.* Joint Bone Spine, 2010. **77**(4): p. 319-324.
75. Firestein, G.S. and I.B. McInnes, *Immunopathogenesis of rheumatoid arthritis.* Immunity, 2017. **46**(2): p. 183-196.
76. Frisell, T., S. Saevarsdottir, and J. Askling, *Family history of rheumatoid arthritis: an old concept with new developments.* Nature Reviews Rheumatology, 2016. **12**(6): p. 335-343.
77. MacGregor, A.J., et al., *Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2000. **43**(1): p. 30-37.
78. Terao, C., et al., *A twin study of rheumatoid arthritis in the Japanese population.* Modern rheumatology, 2016. **26**(5): p. 685-689.
79. Messemaker, T.C., T.W. Huizinga, and F. Kurreeman, *Immunogenetics of rheumatoid arthritis: understanding functional implications.* Journal of autoimmunity, 2015. **64**: p. 74-81.
80. Hensvold, A.H., et al., *Environmental and genetic factors in the development of anticitrullinated protein antibodies (ACPAs) and ACPA-positive rheumatoid arthritis: an epidemiological investigation in twins.* Annals of the rheumatic diseases, 2015. **74**(2): p. 375-380.
81. Van Steenbergen, H., T. Huizinga, and A. Van Der Helm-Van Mil, *The preclinical phase of rheumatoid arthritis: what is acknowledged and what needs to be assessed.* Arthritis Rheum, 2013. **65**(9): p. 2219-32.
82. Frisell, T., et al., *Familial risks and heritability of rheumatoid arthritis: role of rheumatoid factor/anti–citrullinated protein antibody status, number and type of affected relatives, sex, and age.* Arthritis & Rheumatism, 2013. **65**(11): p. 2773-2782.
83. Raychaudhuri, S., et al., *Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis.* Nature genetics, 2012. **44**(3): p. 291-296.
84. Svendsen, A.J., et al., *On the origin of rheumatoid arthritis: the impact of environment and genes—a population based twin study.* PloS one, 2013. **8**(2): p. e57304.

85. Viatte, S., et al., *Association of HLA-DRB1 haplotypes with rheumatoid arthritis severity, mortality, and treatment response.* Jama, 2015. **313**(16): p. 1645-1656.
86. Okada, Y., et al., *Genetics of rheumatoid arthritis contributes to biology and drug discovery.* Nature, 2014. **506**(7488): p. 376-381.
87. Testa, D., et al., *One year in review 2021: pathogenesis of rheumatoid arthritis.* Clin Exp Rheumatol, 2021. **39**(3): p. 445-52.
88. Eyre, S., et al., *High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis.* Nature genetics, 2012. **44**(12): p. 1336-1340.
89. Yarwood, A., T.W. Huizinga, and J. Worthington, *The genetics of rheumatoid arthritis: risk and protection in different stages of the evolution of RA.* Rheumatology, 2016. **55**(2): p. 199-209.
90. Begovich, A.B., et al., *A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis.* The American Journal of Human Genetics, 2004. **75**(2): p. 330-337.
91. Lee, Y.H., et al., *The association between the PTPN22 C1858T polymorphism and rheumatoid arthritis: a meta-analysis update.* Molecular biology reports, 2012. **39**: p. 3453-3460.
92. Lee, Y.H., et al., *Association between the CTLA-4+ 49 A/G polymorphism and susceptibility to rheumatoid arthritis: a meta-analysis.* Molecular biology reports, 2012. **39**: p. 5599-5605.
93. Zhou, C., et al., *Association between CTLA-4 gene polymorphism and risk of rheumatoid arthritis: a meta-analysis.* Aging (Albany NY), 2021. **13**(15): p. 19397.
94. Korman, B.D., et al., *STAT4: genetics, mechanisms, and implications for autoimmunity.* Current allergy and asthma reports, 2008. **8**(5): p. 398-403.
95. Remmers, E.F., et al., *STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus.* New England Journal of Medicine, 2007. **357**(10): p. 977-986.
96. van Boheemen, L. and D. van Schaardenburg, *Predicting rheumatoid arthritis in at-risk individuals.* Clinical therapeutics, 2019. **41**(7): p. 1286-1298.
97. Hoovestol, R.A. and T.R. Mikuls, *Environmental exposures and rheumatoid arthritis risk.* Current rheumatology reports, 2011. **13**: p. 431-439.
98. Damgaard, D., et al., *Smoking is associated with increased levels of extracellular peptidylarginine deiminase 2 (PAD2) in the lungs.* Clin Exp Rheumatol, 2015. **33**(3): p. 405-408.
99. Regueiro, C., et al., *A predominant involvement of the triple seropositive patients and others with rheumatoid factor in the association of smoking with rheumatoid arthritis.* Scientific reports, 2020. **10**(1): p. 3355.
100. De Pablo, P., et al., *Periodontitis in systemic rheumatic diseases.* Nature Reviews Rheumatology, 2009. **5**(4): p. 218-224.
101. Wegner, N., et al., *Peptidylarginine deiminase from Porphyromonas gingivalis citrullinates human fibrinogen and α-enolase: Implications for autoimmunity in rheumatoid arthritis.* Arthritis & Rheumatism, 2010. **62**(9): p. 2662-2672.
102. Chang, K.-H., et al., *Air pollution exposure increases the risk of rheumatoid arthritis: a longitudinal and nationwide study.* Environment international, 2016. **94**: p. 495-499.

103. Hart, J.E., et al., *Ambient air pollution exposures and risk of rheumatoid arthritis: results from the Swedish EIRA case–control study.* Annals of the rheumatic diseases, 2013. **72**(6): p. 888-894.
104. De Roos, A.J., et al., *Proximity to traffic, ambient air pollution, and community noise in relation to incident rheumatoid arthritis.* Environmental health perspectives, 2014. **122**(10): p. 1075-1080.
105. Di, D., et al. *Long-term exposure to outdoor air pollution and the risk of development of rheumatoid arthritis: a systematic review and meta-analysis*. in *Seminars in arthritis and rheumatism*. 2020. Elsevier.
106. Klareskog, L., P.K. Gregersen, and T.W. Huizinga, *Prevention of autoimmune rheumatic disease: state of the art and future perspectives.* Annals of the rheumatic diseases, 2010. **69**(12): p. 2062-2066.
107. Berglin, E., et al., *A combination of autoantibodies to cyclic citrullinated peptide (CCP) and HLA-DRB1 locus antigens is strongly associated with future onset of rheumatoid arthritis.* Arthritis Res Ther, 2004. **6**(4): p. 1-6.
108. Johansson, M., et al., *PTPN22 polymorphism and anti-cyclic citrullinated peptide antibodies in combination strongly predicts future onset of rheumatoid arthritis and has a specificity of 100% for the disease.* Arthritis research & therapy, 2005. **8**: p. 1-6.
109. Makrygiannakis, D., et al., *Citrullination is an inflammation-dependent process.* Annals of the rheumatic diseases, 2006. **65**(9): p. 1219-1222.
110. Spengler, J., et al., *Release of active peptidyl arginine deiminases by neutrophils can explain production of extracellular citrullinated autoantigens in rheumatoid arthritis synovial fluid.* Arthritis & rheumatology, 2015. **67**(12): p. 3135-3145.
111. Neeli, I., S.N. Khan, and M. Radic, *Histone deimination as a response to inflammatory stimuli in neutrophils.* The Journal of Immunology, 2008. **180**(3): p. 1895-1902.
112. Pratesi, F., et al., *Antibodies from patients with rheumatoid arthritis target citrullinated histone 4 contained in neutrophils extracellular traps.* Annals of the rheumatic diseases, 2014. **73**(7): p. 1414-1422.
113. Khandpur, R., et al., *NETs are a source of citrullinated autoantigens and stimulate inflammatory responses in rheumatoid arthritis.* Science translational medicine, 2013. **5**(178): p. 178ra40-178ra40.
114. Brinkmann, V., et al., *Neutrophil extracellular traps kill bacteria.* science, 2004. **303**(5663): p. 1532-1535.
115. Garcia-Romo, G.S., et al., *Netting neutrophils are major inducers of type I IFN production in pediatric systemic lupus erythematosus.* Science translational medicine, 2011. **3**(73): p. 73ra20-73ra20.
116. Gupta, A.K., et al., *Activated endothelial cells induce neutrophil extracellular traps and are susceptible to NETosis-mediated cell death.* FEBS letters, 2010. **584**(14): p. 3193-3197.
117. Lood, C., et al., *Neutrophil extracellular traps enriched in oxidized mitochondrial DNA are interferogenic and contribute to lupus-like disease.* Nature medicine, 2016. **22**(2): p. 146-153.
118. Linn-Rasker, S.P., et al., *Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients who carry HLA-DRB1 shared epitope alleles.* Annals of the rheumatic diseases, 2006. **65**(3): p. 366-371.
119. Meng, W., et al., *DNA methylation mediates genotype and smoking interaction in the development of anti-citrullinated peptide antibody-*

positive rheumatoid arthritis. Arthritis research & therapy, 2017. **19**(1): p. 1-10.
120. Klareskog, L., et al., *A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA–DR (shared epitope)–restricted immune reactions to autoantigens modified by citrullination.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2006. **54**(1): p. 38-46.
121. Johannsen, A., C. Susin, and A. Gustafsson, *Smoking and inflammation: evidence for a synergistic role in chronic disease.* Periodontology 2000, 2014. **64**(1): p. 111-126.
122. Arnson, Y., Y. Shoenfeld, and H. Amital, *Effects of tobacco smoke on immunity, inflammation and autoimmunity.* Journal of autoimmunity, 2010. **34**(3): p. J258-J265.
123. Deane, K.D., et al., *Genetic and environmental risk factors for rheumatoid arthritis.* Best practice & research Clinical rheumatology, 2017. **31**(1): p. 3-18.
124. Demoruelle, M.K., K.D. Deane, and V.M. Holers, *When and where does inflammation begin in rheumatoid arthritis?* Current opinion in rheumatology, 2014. **26**(1): p. 64.
125. Holers, V.M., et al., *Rheumatoid arthritis and the mucosal origins hypothesis: protection turns to destruction.* Nature Reviews Rheumatology, 2018. **14**(9): p. 542-557.
126. Morris, A., et al., *Comparison of the respiratory microbiome in healthy nonsmokers and smokers.* American journal of respiratory and critical care medicine, 2013. **187**(10): p. 1067-1075.
127. Quirke, A.M., et al., *Bronchiectasis is a model for chronic bacterial infection inducing autoimmunity in rheumatoid arthritis.* Arthritis & Rheumatology, 2015. **67**(9): p. 2335-2342.
128. Demoruelle, M.K., et al., *The lung microbiome differs in asymptomatic subjects at elevated risk of future rheumatoid arthritis compared with healthy control subjects.* Annals of the American Thoracic Society, 2014. **11**(Supplement 1): p. S74-S74.
129. Scher, J.U., et al., *The lung microbiota in early rheumatoid arthritis and autoimmunity.* Microbiome, 2016. **4**(1): p. 1-10.
130. Scher, J.U., et al., *Periodontal disease and the oral microbiota in new-onset rheumatoid arthritis.* Arthritis & Rheumatism, 2012. **64**(10): p. 3083-3094.
131. Konig, M.F., et al., *Aggregatibacter actinomycetemcomitans–induced hypercitrullination links periodontal infection to autoimmunity in rheumatoid arthritis.* Science translational medicine, 2016. **8**(369): p. 369ra176-369ra176.
132. Laugisch, O., et al., *Citrullination in the periodontium—a possible link between periodontitis and rheumatoid arthritis.* Clinical oral investigations, 2016. **20**: p. 675-683.
133. Kharlamova, N., et al., *Antibodies to Porphyromonas gingivalis indicate interaction between oral infection, smoking, and risk genes in rheumatoid arthritis etiology.* Arthritis & rheumatology, 2016. **68**(3): p. 604-613.
134. Konig, M.F., et al., *Defining the role of Porphyromonas gingivalis peptidylarginine deiminase (PPAD) in rheumatoid arthritis through the study of PPAD biology.* Annals of the rheumatic diseases, 2015. **74**(11): p. 2054-2061.

135. Harvey, G., et al., *Expression of peptidylarginine deiminase-2 and-4, citrullinated proteins and anti-citrullinated protein antibodies in human gingiva.* Journal of periodontal research, 2013. **48**(2): p. 252-261.
136. Christen, U., et al., *Viral triggers for autoimmunity: is the 'glass of molecular mimicry'half full or half empty?* Journal of autoimmunity, 2010. **34**(1): p. 38-44.
137. Oldstone, M.B., *Molecular mimicry and immune-mediated diseases.* The FASEB Journal, 1998. **12**(13): p. 1255.
138. Wucherpfennig, K.W., *Mechanisms for the induction of autoimmunity by infectious agents.* The Journal of clinical investigation, 2001. **108**(8): p. 1097-1104.
139. Demmer, R.T., et al., *Periodontal disease, tooth loss and incident rheumatoid arthritis: results from the First National Health and Nutrition Examination Survey and its epidemiological follow-up study.* Journal of clinical periodontology, 2011. **38**(11): p. 998-1006.
140. Hitchon, C.A., et al., *Antibodies to Porphyromonas gingivalis are associated with anticitrullinated protein antibodies in patients with rheumatoid arthritis and their relatives.* The Journal of rheumatology, 2010. **37**(6): p. 1105-1112.
141. Mikuls, T.R., et al., *Periodontitis and Porphyromonas gingivalis in patients with rheumatoid arthritis.* Arthritis & rheumatology, 2014. **66**(5): p. 1090-1100.
142. Pischon, N., et al., *Association among rheumatoid arthritis, oral hygiene, and periodontitis.* Journal of periodontology, 2008. **79**(6): p. 979-986.
143. Crowson, C.S., et al., *Rheumatoid arthritis and cardiovascular disease.* American heart journal, 2013. **166**(4): p. 622-628. e1.
144. Prieto-Alhambra, D., et al., *An increased rate of falling leads to a rise in fracture risk in postmenopausal women with self-reported osteoarthritis: a prospective multinational cohort study (GLOW).* Annals of the rheumatic diseases, 2013. **72**(6): p. 911-917.
145. Doran, M.F., et al., *Frequency of infection in patients with rheumatoid arthritis compared with controls: a population-based study.* Arthritis & Rheumatism, 2002. **46**(9): p. 2287-2293.
146. Cojocaru, M., et al., *Extra-articular manifestations in rheumatoid arthritis.* Maedica, 2010. **5**(4): p. 286.
147. Liao, K.P., L. Alfredsson, and E.W. Karlson, *Environmental influences on risk for rheumatoid arthritis.* Current opinion in rheumatology, 2009. **21**(3): p. 279-283.
148. Brzustewicz, E. and E. Bryl, *The role of cytokines in the pathogenesis of rheumatoid arthritis–Practical and potential application of cytokines as biomarkers and targets of personalized therapy.* Cytokine, 2015. **76**(2): p. 527-536.
149. Müller-Ladner, U. and T. Pap, *Pathogenesis of RA: more than just immune cells.* Zeitschrift für Rheumatologie, 2005. **64**: p. 396-401.
150. Mateen, S., et al., *Understanding the role of cytokines in the pathogenesis of rheumatoid arthritis.* Clinica chimica acta, 2016. **455**: p. 161-171.
151. Alam, J., I. Jantan, and S.N.A. Bukhari, *Rheumatoid arthritis: Recent advances on its etiology, role of cytokines and pharmacotherapy.* Biomedicine & Pharmacotherapy, 2017. **92**: p. 615-633.
152. Winchester, R., *The molecular basis of susceptibility to rheumatoid arthritis.* Advances in immunology, 1994. **56**: p. 389-466.

153. Gregersen, P.K., J. Silver, and R.J. Winchester, *The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis.* Arthritis & Rheumatism, 1987. **30**(11): p. 1205-1213.
154. Conrad, K., et al., *Profiling of rheumatoid arthritis associated autoantibodies.* Autoimmunity reviews, 2010. **9**(6): p. 431-435.
155. Vander Cruyssen, B., et al., *Anti-citrullinated protein/peptide antibodies (ACPA) in rheumatoid arthritis: specificity and relation with rheumatoid factor.* Autoimmunity reviews, 2005. **4**(7): p. 468-474.
156. Derksen, V., T. Huizinga, and D. Van Der Woude. *The role of autoantibodies in the pathophysiology of rheumatoid arthritis*. in *Seminars in immunopathology*. 2017. Springer.
157. Song, Y. and E. Kang, *Autoantibodies in rheumatoid arthritis: rheumatoid factors and anticitrullinated protein antibodies.* QJM: An International Journal of Medicine, 2010. **103**(3): p. 139-146.
158. Feldmann, M., F.M. Brennan, and R.N. Maini, *Role of cytokines in rheumatoid arthritis.* Annual review of immunology, 1996. **14**(1): p. 397-440.
159. Goëb, V., et al., *Potential role of arthroscopy in the management of inflammatory arthritis.* Clinical and Experimental Rheumatology-Incl Supplements, 2012. **30**(3): p. 429.
160. Panayi, G., J. Lanchbury, and G. Kingsley, *The importance of the T cell in initiating and maintaining the chronic synovitis of rheumatoid arthritis.* Arthritis and rheumatism, 1992. **35**(7): p. 729-735.
161. Salmon, M. and J. Hill Gaston, *The role of T-lymphocytes in rheumatoid arthritis.* British medical bulletin, 1995. **51**(2): p. 332-345.
162. Lanchbury, J.S. and C. Pitzalis, *Cellular immune mechanisms in rheumatoid arthritis and other inflammatory arthritides.* Current opinion in immunology, 1993. **5**(6): p. 918-924.
163. Fiocco, U., et al., *Co-stimulatory modulation in rheumatoid arthritis: the role of (CTLA4-Ig) abatacept.* Autoimmunity reviews, 2008. **8**(1): p. 76-82.
164. Weyand, C.M. and J.J. Goronzy, *T-cell-targeted therapies in rheumatoid arthritis.* Nature Clinical Practice Rheumatology, 2006. **2**(4): p. 201-210.
165. Bakakos, P., et al., *Simultaneous analysis of T cell clonality and cytokine production in rheumatoid arthritis using three-colour flow cytometry.* Clinical & Experimental Immunology, 2002. **129**(2): p. 370-378.
166. Musters, A., et al., *In rheumatoid arthritis, synovitis at different inflammatory sites is dominated by shared but patient-specific T cell clones.* The Journal of Immunology, 2018. **201**(2): p. 417-422.
167. Majka, D.S., et al., *Duration of preclinical rheumatoid arthritis-related autoantibody positivity increases in subjects with older age at time of disease diagnosis.* Annals of the rheumatic diseases, 2008. **67**(6): p. 801-807.
168. Ramos-Remus, C., et al., *Autoantibodies in prediction of the development of rheumatoid arthritis among healthy relatives of patients with the disease.* Arthritis & rheumatology, 2015. **67**(11): p. 2837-2844.
169. Deane, K.D. and V.M. Holers, *Rheumatoid arthritis pathogenesis, prediction, and prevention: an emerging paradigm shift.* Arthritis & Rheumatology, 2021. **73**(2): p. 181-193.

170. Gan, R.W., et al., *Anti-carbamylated protein antibodies are present prior to rheumatoid arthritis and are associated with its future diagnosis.* The Journal of rheumatology, 2015. **42**(4): p. 572-579.
171. Kolfenbach, J.R., et al., *Autoimmunity to peptidyl arginine deiminase type 4 precedes clinical onset of rheumatoid arthritis.* Arthritis & Rheumatism, 2010. **62**(9): p. 2633-2639.
172. Sokolove, J., et al., *Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis.* PloS one, 2012. **7**(5): p. e35296.
173. Kokkonen, H., et al., *Antibodies of IgG, IgA and IgM isotypes against cyclic citrullinated peptide precede the development of rheumatoid arthritis.* Arthritis research & therapy, 2011. **13**: p. 1-10.
174. Kokkonen, H., et al., *Associations of antibodies against citrullinated peptides with human leukocyte antigen-shared epitope and smoking prior to the development of rheumatoid arthritis.* Arthritis research & therapy, 2015. **17**(1): p. 1-15.
175. Kelmenson, L.B., et al., *Timing of elevations of autoantibody isotypes prior to diagnosis of rheumatoid arthritis.* Arthritis & Rheumatology, 2020. **72**(2): p. 251-261.
176. Novella-Navarro, M., et al., *Risk factors for developing rheumatoid arthritis in patients with undifferentiated arthritis and inflammatory arthralgia.* Frontiers in Medicine, 2021. **8**: p. 668898.
177. Kokkonen, H., et al., *Up-regulation of cytokines and chemokines predates the onset of rheumatoid arthritis.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2010. **62**(2): p. 383-391.
178. Deane, K.D., et al., *The number of elevated cytokines and chemokines in preclinical seropositive rheumatoid arthritis predicts time to diagnosis in an age-dependent manner.* Arthritis & Rheumatism, 2010. **62**(11): p. 3161-3172.
179. Patakas, A., et al., *Breach of self tolerance in rheumatoid arthritis: a role for Th17 effector T cells?* Annals of the Rheumatic Diseases, 2011. **70**(Suppl 2): p. A50-A50.
180. O'Neil, L.J. and M.J. Kaplan, *Neutrophils in rheumatoid arthritis: breaking immune tolerance and fueling disease.* Trends in molecular medicine, 2019. **25**(3): p. 215-227.
181. Ioan-Facsinay, A., et al., *Marked differences in fine specificity and isotype usage of the anti–citrullinated protein antibody in health and disease.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2008. **58**(10): p. 3000-3008.
182. Van Venrooij, W.J., J.J. Van Beers, and G.J. Pruijn, *Anti-CCP antibodies: the past, the present and the future.* Nature Reviews Rheumatology, 2011. **7**(7): p. 391-398.
183. Deane, K.D., J.M. Norris, and V.M. Holers, *Preclinical rheumatoid arthritis: identification, evaluation, and future directions for investigation.* Rheumatic Disease Clinics, 2010. **36**(2): p. 213-241.
184. De Hair, M., et al., *Features of the synovium of individuals at risk of developing rheumatoid arthritis: implications for understanding preclinical rheumatoid arthritis.* Arthritis & rheumatology, 2014. **66**(3): p. 513-522.
185. van de Stadt, L.A., et al., *A prediction rule for the development of arthritis in seropositive arthralgia patients.* Annals of the rheumatic diseases, 2013. **72**(12): p. 1920-1926.

186. Raposo, B., et al., *T cells specific for post-translational modifications escape intrathymic tolerance induction.* Nature communications, 2018. **9**(1): p. 353.
187. Weyand, C.M. and J.J. Goronzy, *The immunology of rheumatoid arthritis.* Nature immunology, 2021. **22**(1): p. 10-18.
188. Tracy, A., C.D. Buckley, and K. Raza. *Pre-symptomatic autoimmunity in rheumatoid arthritis: when does the disease start?* in *Seminars in immunopathology*. 2017. Springer.
189. van de Stadt, L.A., et al., *Development of the anti–citrullinated protein antibody repertoire prior to the onset of rheumatoid arthritis.* Arthritis & Rheumatism, 2011. **63**(11): p. 3226-3233.
190. Suwannalai, P., et al., *Avidity maturation of anti–citrullinated protein antibodies in rheumatoid arthritis.* Arthritis & Rheumatism, 2012. **64**(5): p. 1323-1328.
191. Kissel, T., et al., *Antibodies and B cells recognising citrullinated proteins display a broad cross-reactivity towards other post-translational modifications.* Annals of the Rheumatic Diseases, 2020. **79**(4): p. 472-480.
192. Castro-Sánchez, P. and P. Roda-Navarro, *Physiology and pathology of autoimmune diseases: role of CD4+ T cells in rheumatoid arthritis*, in *Physiology and Pathology of Immunology*. 2017, IntechOpen.
193. McInnes, I.B. and G. Schett, *The pathogenesis of rheumatoid arthritis.* New England Journal of Medicine, 2011. **365**(23): p. 2205-2219.
194. Malmström, V., A.I. Catrina, and L. Klareskog, *The immunopathogenesis of seropositive rheumatoid arthritis: from triggering to targeting.* Nature Reviews Immunology, 2017. **17**(1): p. 60-75.
195. Scally, S.W., et al., *A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis.* Journal of Experimental Medicine, 2013. **210**(12): p. 2569-2582.
196. Malmström, V. and C. Grönwall, *The parallel worlds of ACPA-positive and RF-positive B cells.* Nature Reviews Rheumatology, 2018. **14**(11): p. 626-628.
197. Mahendra, A., et al., *Beyond autoantibodies: biologic roles of human autoreactive B cells in rheumatoid arthritis revealed by RNA-sequencing.* Arthritis & Rheumatology, 2019. **71**(4): p. 529-541.
198. He, J., J. Ju, and X. Wang, *The current status of anti-citrullinated protein antibodies and citrullinated protein-reactive B cells in the pathogenesis of rheumatoid arthritis.* Molecular Biology Reports, 2022: p. 1-11.
199. Bonilla, F.A. and H.C. Oettgen, *Adaptive immunity.* Journal of Allergy and Clinical Immunology, 2010. **125**(2): p. S33-S40.
200. Appleman, L.J., et al., *CD28 costimulation mediates T cell expansion via IL-2-independent and IL-2-dependent regulation of cell cycle progression.* The Journal of Immunology, 2000. **164**(1): p. 144-151.
201. McNally, A., et al., *CD4+ CD25+ regulatory T cells control CD8+ T-cell effector differentiation by modulating IL-2 homeostasis.* Proceedings of the National Academy of Sciences, 2011. **108**(18): p. 7529-7534.
202. Luo, P., et al., *Immunomodulatory role of T helper cells in rheumatoid arthritis: A comprehensive research review.* Bone & Joint Research, 2022. **11**(7): p. 426-438.
203. Podojil, J.R. and S.D. Miller, *Molecular mechanisms of T-cell receptor and costimulatory molecule ligation/blockade in autoimmune disease therapy.* Immunological reviews, 2009. **229**(1): p. 337-355.

204. Meednu, N., et al., *Production of RANKL by memory B cells: a link between B cells and bone erosion in rheumatoid arthritis.* Arthritis & rheumatology, 2016. **68**(4): p. 805-816.
205. Van der Graaff, W., et al., *Quantitation of interferon gamma-and interleukin-4-producing T cells in synovial fluid and peripheral blood of arthritis patients.* Rheumatology (Oxford, England), 1999. **38**(3): p. 214-220.
206. Berner, B., et al., *Analysis of Th1 and Th2 cytokines expressing CD4+ and CD8+ T cells in rheumatoid arthritis by flow cytometry.* The Journal of rheumatology, 2000. **27**(5): p. 1128-1135.
207. Zhang, R., et al., *Th1-like treg cells are increased but deficient in function in rheumatoid arthritis.* Frontiers in Immunology, 2022. **13**: p. 863753.
208. Julé, A.M., et al., *Th1 polarization defines the synovial fluid T cell compartment in oligoarticular juvenile idiopathic arthritis.* JCI insight, 2021. **6**(18).
209. Haddad, A., J. Bienvenu, and P. Miossec, *Increased production of a Th2 cytokine profile by activated whole blood cells from rheumatoid arthritis patients.* Journal of clinical immunology, 1998. **18**: p. 399-403.
210. Masoumi, M., et al., *Role of T cells in the pathogenesis of rheumatoid arthritis: focus on immunometabolism dysfunctions.* Inflammation, 2023. **46**(1): p. 88-102.
211. Alunno, A., et al., *Altered immunoregulation in rheumatoid arthritis: the role of regulatory T cells and proinflammatory Th17 cells and therapeutic implications.* Mediators of inflammation, 2015. **2015**.
212. Al-Saadany, H.M., et al., *Th-17 cells and serum IL-17 in rheumatoid arthritis patients: Correlation with disease activity and severity.* The Egyptian Rheumatologist, 2016. **38**(1): p. 1-7.
213. Cooles, F.A., J.D. Isaacs, and A.E. Anderson, *Treg cells in rheumatoid arthritis: an update.* Current rheumatology reports, 2013. **15**: p. 1-9.
214. Boissier, M.-C., et al., *Regulatory T cells (Treg) in rheumatoid arthritis.* Joint Bone Spine, 2009. **76**(1): p. 10-14.
215. Morita, T., et al., *The proportion of regulatory T cells in patients with rheumatoid arthritis: a meta-analysis.* PloS one, 2016. **11**(9): p. e0162306.
216. Rapetti, L., et al., *B cell resistance to Fas-mediated apoptosis contributes to their ineffective control by regulatory T cells in rheumatoid arthritis.* Annals of the rheumatic diseases, 2015. **74**(1): p. 294-302.
217. Al-Zifzaf, D.S., et al., *FoxP3+ T regulatory cells in Rheumatoid arthritis and the imbalance of the Treg/TH17 cytokine axis.* The Egyptian Rheumatologist, 2015. **37**(1): p. 7-15.
218. Schulze-Koops, H. and J.R. Kalden, *The balance of Th1/Th2 cytokines in rheumatoid arthritis.* Best practice & research Clinical rheumatology, 2001. **15**(5): p. 677-691.
219. Ciccia, F., et al., *Potential involvement of IL-9 and Th9 cells in the pathogenesis of rheumatoid arthritis.* Rheumatology, 2015. **54**(12): p. 2264-2272.
220. Chowdhury, K., et al., *Synovial IL-9 facilitates neutrophil survival, function and differentiation of Th17 cells in rheumatoid arthritis.* Arthritis research & therapy, 2018. **20**: p. 1-12.
221. Cope, A.P., H. Schulze-Koops, and M. Aringer, *The central role of T cells in rheumatoid arthritis.* Clinical and experimental rheumatology, 2007. **25**(5): p. S4.

222. Rao, D.A., *T cells that help B cells in chronically inflamed tissues.* Frontiers in immunology, 2018. **9**: p. 402904.
223. Rao, D.A., et al., *Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis.* Nature, 2017. **542**(7639): p. 110-114.
224. Salmon, M., et al., *The progressive differentiation of primed T cells is associated with an increasing susceptibility to apoptosis.* European journal of immunology, 1994. **24**(4): p. 892-899.
225. Sallusto, F. and A. Lanzavecchia, *Exploring pathways for memory T cell generation.* The Journal of clinical investigation, 2001. **108**(6): p. 805-806.
226. Sallusto, F., et al., *Two subsets of memory T lymphocytes with distinct homing potentials and effector functions.* Nature, 1999. **401**(6754): p. 708-712.
227. Muñoz-Espín, D. and M. Serrano, *Cellular senescence: from physiology to pathology.* Nature reviews Molecular cell biology, 2014. **15**(7): p. 482-496.
228. Pilling, D., et al., *CD4+ CD45RA+ T cells from adults respond to recall antigens after CD28 ligation.* International immunology, 1996. **8**(11): p. 1737-1742.
229. Arlettaz, L., et al., *CD45 isoform phenotypes of human T cells: CD4+ CD45RA–RO+ memory T cells re-acquire CD45RA without losing CD45RO.* European journal of immunology, 1999. **29**(12): p. 3987-3994.
230. Douek, D.C., et al., *Changes in thymic function with age and during the treatment of HIV infection.* Nature, 1998. **396**(6712): p. 690-695.
231. Poulin, J.-F., et al., *Direct evidence for thymic function in adult humans.* The Journal of experimental medicine, 1999. **190**(4): p. 479-486.
232. Hazenberg, M.D., et al., *Increased cell division but not thymic dysfunction rapidly affects the T-cell receptor excision circle content of the naive T cell population in HIV-1 infection.* Nature medicine, 2000. **6**(9): p. 1036-1042.
233. Hazenberg, M.D., et al., *T-cell receptor excision circle and T-cell dynamics after allogeneic stem cell transplantation are related to clinical events.* Blood, The Journal of the American Society of Hematology, 2002. **99**(9): p. 3449-3453.
234. Kaech, S.M. and R. Ahmed, *Memory CD8+ T cell differentiation: initial antigen encounter triggers a developmental program in naive cells.* Nature immunology, 2001. **2**(5): p. 415-422.
235. Aarvak, T. and J.B. Natvig, *Cell-cell interactions in synovitis: antigen presenting cells and T cell interaction in rheumatoid arthritis.* Arthritis Research & Therapy, 2000. **3**: p. 1-5.
236. Yu, M.B. and W.H. Langridge, *The function of myeloid dendritic cells in rheumatoid arthritis.* Rheumatology International, 2017. **37**: p. 1043-1051.
237. Fox, D.A., et al., *Cell-cell interactions in rheumatoid arthritis synovium.* Rheumatic Disease Clinics, 2010. **36**(2): p. 311-323.
238. Rao, D.A., *T cells that help B cells in chronically inflamed tissues.* Frontiers in immunology, 2018. **9**: p. 1924.
239. Sagaert, X. and C. De Wolf-Peeters, *Classification of B-cells according to their differentiation status, their micro-anatomical localisation and their developmental lineage.* Immunology letters, 2003. **90**(2-3): p. 179-186.
240. Moura, R.A., L. Graca, and J.E. Fonseca, *To B or not to B the conductor of rheumatoid arthritis orchestra.* Clinical reviews in allergy & immunology, 2012. **43**: p. 281-291.

241. Panayi, G., *B cells: a fundamental role in the pathogenesis of rheumatoid arthritis?* Rheumatology, 2005. **44**(suppl_2): p. ii3-ii7.
242. Bugatti, S., et al., *B cells in rheumatoid arthritis: from pathogenic players to disease biomarkers.* BioMed research international, 2014. **2014**.
243. Rubin, S.J., M.S. Bloom, and W.H. Robinson, *B cell checkpoints in autoimmune rheumatic diseases.* Nature Reviews Rheumatology, 2019. **15**(5): p. 303-315.
244. Kramm, H., et al., *Successful therapy of rheumatoid arthritis with rituximab: renewed interest in the role of B cells in the pathogenesis of rheumatoid arthritis.* JCR: Journal of Clinical Rheumatology, 2004. **10**(1): p. 28-32.
245. Mikuls, T.R., et al., *Association of rheumatoid arthritis treatment response and disease duration with declines in serum levels of IgM rheumatoid factor and anti–cyclic citrullinated peptide antibody.* Arthritis & Rheumatism, 2004. **50**(12): p. 3776-3782.
246. Vos, K., et al., *Early effects of rituximab on the synovial cell infiltrate in patients with rheumatoid arthritis.* Arthritis & Rheumatism, 2007. **56**(3): p. 772-778.
247. Walsh, C., et al., *Decreased CD20 expression in rheumatoid arthritis synovium following 8 weeks of rituximab therapy.* Clinical & Experimental Rheumatology, 2008. **26**(4): p. 656.
248. Aho, K., et al., *Rheumatoid factors antedating clinical rheumatoid arthritis.* The Journal of rheumatology, 1991. **18**(9): p. 1282-1284.
249. Puente, A.D., et al., *The incidence of rheumatoid arthritis is predicted by rheumatoid factor titer in a longitudinal population study.* Arthritis & Rheumatism, 1988. **31**(10): p. 1239-1244.
250. Halldorsdottir, H., et al., *A prospective study on the incidence of rheumatoid arthritis among people with persistent increase of rheumatoid factor.* Annals of the rheumatic diseases, 2000. **59**(2): p. 149-151.
251. Nielen, M.M., et al., *Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2004. **50**(2): p. 380-386.
252. Hua, Z. and B. Hou, *TLR signaling in B-cell development and activation.* Cellular & molecular immunology, 2013. **10**(2): p. 103-106.
253. O'Garra, A., et al., *Production of cytokines by mouse B cells: B lymphomas and normal B cells produce interleukin 10.* International immunology, 1990. **2**(9): p. 821-832.
254. Pistoia, V., *Production of cytokines by human B cells in health and disease.* Immunology today, 1997. **18**(7): p. 343-350.
255. Harris, D.P., et al., *Reciprocal regulation of polarized cytokine production by effector B and T cells.* Nature immunology, 2000. **1**(6): p. 475-482.
256. Snider, D. and D. Segal, *Efficiency of antigen presentation after antigen targeting to surface IgD, IgM, MHC, Fc gamma RII, and B220 molecules on murine splenic B cells.* Journal of immunology (Baltimore, Md.: 1950), 1989. **143**(1): p. 59-65.
257. Batista, F.D. and M.S. Neuberger, *Affinity dependence of the B cell response to antigen: a threshold, a ceiling, and the importance of off-rate.* Immunity, 1998. **8**(6): p. 751-759.
258. Roth, R., T. Nakamura, and M.J. Mamula, *B7 costimulation and autoantigen specificity enable B cells to activate autoreactive T cells.*

Journal of immunology (Baltimore, Md.: 1950), 1996. **157**(7): p. 2924-2931.
259. Rodríguez-Pinto, D., *B cells as antigen presenting cells.* Cellular immunology, 2005. **238**(2): p. 67-75.
260. Lehmann, P.V., et al., *Determinant spreading and the dynamics of the autoimmune T-cell repertoire.* Immunology today, 1993. **14**(5): p. 203-208.
261. Lemoine, S., et al., *Human T cells induce their own regulation through activation of B cells.* Journal of Autoimmunity, 2011. **36**(3-4): p. 228-238.
262. Frey, S., et al., *The novel cytokine interleukin-36α is expressed in psoriatic and rheumatoid arthritis synovium.* Annals of the rheumatic diseases, 2013. **72**(9): p. 1569-1574.
263. Kimura, A. and T. Kishimoto, *IL-6: regulator of Treg/Th17 balance.* European journal of immunology, 2010. **40**(7): p. 1830-1835.
264. Mitsdoerffer, M., et al., *Proinflammatory T helper type 17 cells are effective B-cell helpers.* Proceedings of the National Academy of Sciences, 2010. **107**(32): p. 14292-14297.
265. Schlegel, P.M., et al., *B cells contribute to heterogeneity of IL-17 producing cells in rheumatoid arthritis and healthy controls.* PloS one, 2013. **8**(12): p. e82580.
266. Thiolat, A., et al., *Interleukin-6 receptor blockade enhances CD39+ regulatory T cell development in rheumatoid arthritis and in experimental arthritis.* Arthritis & rheumatology, 2014. **66**(2): p. 273-283.
267. van de Veerdonk, F.L., et al., *The anti-CD20 antibody rituximab reduces the Th17 cell response.* Arthritis & Rheumatism, 2011. **63**(6): p. 1507-1516.
268. Wong, P.K., et al., *Interleukin-6 modulates production of T lymphocyte–derived cytokines in antigen-induced arthritis and drives inflammation-induced osteoclastogenesis.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2006. **54**(1): p. 158-168.
269. Yeo, L., et al., *Cytokine mRNA profiling identifies B cells as a major source of RANKL in rheumatoid arthritis.* Annals of the rheumatic diseases, 2011. **70**(11): p. 2022-2028.
270. Flores-Borja, F., et al., *CD19+ CD24hiCD38hi B cells maintain regulatory T cells while limiting TH1 and TH17 differentiation.* Science translational medicine, 2013. **5**(173): p. 173ra23-173ra23.
271. Blair, P.A., et al., *CD19+ CD24hiCD38hi B cells exhibit regulatory capacity in healthy individuals but are functionally impaired in systemic lupus erythematosus patients.* immunity, 2010. **32**(1): p. 129-140.
272. Ma, L., et al., *Reduced numbers of regulatory B cells are negatively correlated with disease activity in patients with new-onset rheumatoid arthritis.* Clinical rheumatology, 2014. **33**: p. 187-195.
273. Haringman, J.J., et al., *Synovial tissue macrophages: a sensitive biomarker for response to treatment in patients with rheumatoid arthritis.* Annals of the rheumatic diseases, 2005. **64**(6): p. 834-838.
274. Kinne, R.W., et al., *Macrophages in rheumatoid arthritis.* Arthritis Research & Therapy, 2000. **2**(3): p. 1-14.
275. Siouti, E. and E. Andreakos, *The many facets of macrophages in rheumatoid arthritis.* Biochemical pharmacology, 2019. **165**: p. 152-169.
276. Ardura, J.A., et al., *Targeting macrophages: friends or foes in disease?* Frontiers in Pharmacology, 2019. **10**: p. 1255.

277. Edilova, M.I., A. Akram, and A.A. Abdul-Sater, *Innate immunity drives pathogenesis of rheumatoid arthritis.* biomedical journal, 2021. **44**(2): p. 172-182.
278. Tardito, S., et al., *Macrophage M1/M2 polarization and rheumatoid arthritis: a systematic review.* Autoimmunity reviews, 2019. **18**(11): p. 102397.
279. Fukui, S., et al., *M1 and M2 monocytes in rheumatoid arthritis: a contribution of imbalance of M1/M2 monocytes to osteoclastogenesis.* Frontiers in immunology, 2018. **8**: p. 1958.
280. Shigeyama, Y., et al., *Expression of osteoclast differentiation factor in rheumatoid arthritis.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2000. **43**(11): p. 2523-2530.
281. Guerne, P.-A., et al., *Synovium as a source of interleukin 6 in vitro. Contribution to local and systemic manifestations of arthritis.* The Journal of clinical investigation, 1989. **83**(2): p. 585-592.
282. Davignon, J.-L., et al., *Targeting monocytes/macrophages in the treatment of rheumatoid arthritis.* Rheumatology, 2013. **52**(4): p. 590-598.
283. Pap, T., et al., *Fibroblast biology: role of synovial fibroblasts in the pathogenesis of rheumatoid arthritis.* Arthritis Research & Therapy, 2000. **2**(5): p. 1-7.
284. Pap, T., et al., *Are fibroblasts involved in joint destruction?* Annals of the Rheumatic Diseases, 2005. **64**(suppl 4): p. iv52-iv54.
285. Bottini, N. and G.S. Firestein, *Duality of fibroblast-like synoviocytes in RA: passive responders and imprinted aggressors.* Nature Reviews Rheumatology, 2013. **9**(1): p. 24-33.
286. Bustamante, M.F., et al., *Fibroblast-like synoviocyte metabolism in the pathogenesis of rheumatoid arthritis.* Arthritis research & therapy, 2017. **19**(1): p. 1-12.
287. Shegarfi, H., F. Naddafi, and A. Mirshafiey, *Natural killer cells and their role in rheumatoid arthritis: friend or foe?* The Scientific World Journal, 2012. **2012**.
288. Vivier, E., et al., *Functions of natural killer cells.* Nature immunology, 2008. **9**(5): p. 503-510.
289. Tomasello, E., et al., *Mapping of NKp46+ cells in healthy human lymphoid and non-lymphoid tissues.* Frontiers in immunology, 2012. **3**: p. 344.
290. Kiessling, R., E. Klein, and H. Wigzell, *„Natural" killer cells in the mouse. I. Cytotoxic cells with specificity for mouse Moloney leukemia cells. Specificity and distribution according to genotype.* European journal of immunology, 1975. **5**(2): p. 112-117.
291. Dalbeth, N., et al., *CD56bright NK cells are enriched at inflammatory sites and can engage with monocytes in a reciprocal program of activation.* The Journal of Immunology, 2004. **173**(10): p. 6418-6426.
292. Poli, A., et al., *CD56bright natural killer (NK) cells: an important NK cell subset.* Immunology, 2009. **126**(4): p. 458-465.
293. Dalbeth, N. and M.F. Callan, *A subset of natural killer cells is greatly expanded within inflamed joints.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2002. **46**(7): p. 1763-1772.
294. Tak, P.P., et al., *Granzyme-positive cytotoxic cells are specifically increased in early rheumatoid synovial tissue.* Arthritis & Rheumatism, 1994. **37**(12): p. 1735-1743.

295. Khan, S., J.D. Greenberg, and N. Bhardwaj, *Dendritic cells as targets for therapy in rheumatoid arthritis.* Nature Reviews Rheumatology, 2009. **5**(10): p. 566-571.
296. Hilkens, C. and J. Isaacs, *Tolerogenic dendritic cell therapy for rheumatoid arthritis: where are we now?* Clinical & Experimental Immunology, 2013. **172**(2): p. 148-157.
297. Iberg, C.A., A. Jones, and D. Hawiger, *Dendritic cells as inducers of peripheral tolerance.* Trends in immunology, 2017. **38**(11): p. 793-804.
298. Steinman, R.M., et al., *Dendritic cell function in vivo during the steady state: a role in peripheral tolerance.* Annals of the New York Academy of Sciences, 2003. **987**(1): p. 15-25.
299. Deane, K.D. and V.M. Holers, *The natural history of rheumatoid arthritis.* Clinical therapeutics, 2019. **41**(7): p. 1256-1269.
300. Hunt, L. and P. Emery, *Defining populations at risk of rheumatoid arthritis: the first steps to prevention.* Nature Reviews Rheumatology, 2014. **10**(9): p. 521-530.
301. Sahlström, P., et al., *Different hierarchies of anti–modified protein autoantibody reactivities in rheumatoid arthritis.* Arthritis & Rheumatology, 2020. **72**(10): p. 1643-1657.
302. Aletaha, D., et al., *2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative.* Arthritis & rheumatism, 2010. **62**(9): p. 2569-2581.
303. Willemze, A., et al., *The influence of ACPA status and characteristics on the course of RA.* Nature Reviews Rheumatology, 2012. **8**(3): p. 144-152.
304. Singwe-Ngandeu, M., et al., *Diagnostic value of anti-cyclic citrullinated peptides and association with HLA-DRB1 shared epitope alleles in African rheumatoid arthritis patients.* Arthritis research & therapy, 2010. **12**(2): p. 1-7.
305. Nishimura, K., et al., *Meta-analysis: diagnostic accuracy of anti–cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis.* Annals of internal medicine, 2007. **146**(11): p. 797-808.
306. Sun, J., et al., *Diagnostic accuracy of combined tests of anti cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis: a meta-analysis.* Clin Exp Rheumatol, 2014. **32**(1): p. 11-21.
307. Van Venrooij, W.J., J.J. Van Beers, and G.J. Pruijn, *Anti-CCP antibody, a marker for the early detection of rheumatoid arthritis.* Annals of the New York Academy of Sciences, 2008. **1143**(1): p. 268-285.
308. Hunt, L. and P. Emery, *Defining populations at risk of rheumatoid arthritis: the first steps to prevention.* Nat Rev Rheumatol, 2014. **10**(9): p. 521-30.
309. Gerlag, D.M., et al., *EULAR recommendations for terminology and research in individuals at risk of rheumatoid arthritis: report from the Study Group for Risk Factors for Rheumatoid Arthritis.* Annals of the Rheumatic Diseases, 2012. **71**(5): p. 638-641.
310. Kurowska, W., et al., *The role of anti-citrullinated protein antibodies (ACPA) in the pathogenesis of rheumatoid arthritis.* Central European Journal of Immunology, 2017. **42**(4): p. 390-398.
311. Rantapää-Dahlqvist, S., et al., *Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis.* Arthritis & Rheumatism, 2003. **48**(10): p. 2741-2749.

312. Rakieh, C., et al., *Predicting the development of clinical arthritis in anti-CCP positive individuals with non-specific musculoskeletal symptoms: a prospective observational cohort study.* Annals of the rheumatic diseases, 2015. **74**(9): p. 1659-1666.
313. Hunt, L. and M. Buch, *The 'therapeutic window'and treating to target in rheumatoid arthritis.* Clinical medicine, 2013. **13**(4): p. 387.
314. Van Riel, P. and L. Renskers, *The Disease Activity Score (DAS) and the Disease Activity Score using 28 joint counts (DAS28) in the management of rheumatoid arthritis.* Clin Exp Rheumatol, 2016. **34**(5 Suppl 101): p. S40-S44.
315. McWilliams, D.F., et al., *Interpretation of DAS28 and its components in the assessment of inflammatory and non-inflammatory aspects of rheumatoid arthritis.* BMC rheumatology, 2018. **2**(1): p. 1-12.
316. Nice, *Rheumatoid arthritis: The management of rheumatoid arthritis in adults*. 2009, National Institute for Health and Clinical Excellence London.
317. Allaart, C., et al., *Aiming at low disease activity in rheumatoid arthritis with initial combination therapy or initial monotherapy strategies: the BeSt study.* Clinical and experimental rheumatology, 2006. **24**(6 Suppl 43): p. S-77.
318. Fransen, J. and P.L. Van Riel, *The Disease Activity Score and the EULAR response criteria.* Rheumatic Disease Clinics, 2009. **35**(4): p. 745-757.
319. Goekoop-Ruiterman, Y.d., et al., *Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial.* Arthritis & Rheumatism, 2005. **52**(11): p. 3381-3390.
320. Fleischmann, R.M., et al., *DAS28-CRP and DAS28-ESR cut-offs for high disease activity in rheumatoid arthritis are not interchangeable.* RMD open, 2017. **3**(1): p. e000382.
321. Fuchs, H.A., et al., *A simplified twenty-eight–joint quantitative articular index in rheumatoid arthritis.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 1989. **32**(5): p. 531-537.
322. van der Heijde, D.M., et al., *Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score.* Annals of the rheumatic diseases, 1990. **49**(11): p. 916-920.
323. Fransen, J., et al. *Development and validation of the DAS28 using CRP*. in *Annals of the rheumatic diseases*. 2003. BMJ PUBLISHING GROUP BRITISH MED ASSOC HOUSE, TAVISTOCK SQUARE, LONDON WC1H ….
324. Putrik, P., et al., *Variations in criteria regulating treatment with reimbursed biologic DMARDs across European countries. Are differences related to country's wealth?* Annals of the rheumatic diseases, 2014. **73**(11): p. 2010-2021.
325. Hensor, E.M. and P.G. Conaghan, *Time to modify the DAS28 to make it fit for purpose (s) in rheumatoid arthritis?* Expert Review of Clinical Immunology, 2020. **16**(1): p. 1-4.
326. Fransen, J., M. Creemers, and P. Van Riel, *Remission in rheumatoid arthritis: agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria.* Rheumatology, 2004. **43**(10): p. 1252-1255.
327. Cohen, G., et al., *Radiological damage in patients with rheumatoid arthritis on sustained remission.* Annals of the rheumatic diseases, 2007. **66**(3): p. 358-363.

328. Rook, E., *AB0179 Revision of European Medicine Agency's (EMA) Guideline on Clinical Investigation of Medicinal Products Other than Nsaids for Treatment of Rheumatoid Arthritis, 2016*. 2016, BMJ Publishing Group Ltd.
329. McWilliams, D.F., et al., *Predictors of change in bodily pain in early rheumatoid arthritis: an inception cohort study.* Arthritis care & research, 2012. **64**(10): p. 1505-1513.
330. Hensor, E.M., et al., *O12 Validity of a2-component imaging-derived disease activity score (2C-DAS28) for improved assessment of synovitis in early rheumatoid arthritis.* Rheumatology, 2018. **57**(suppl_3): p. key075. 194.
331. Boers, M., et al., *Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0.* Journal of clinical epidemiology, 2014. **67**(7): p. 745-753.
332. Salomon-Escoto, K. and J. Kay, *The "treat to target" approach to rheumatoid arthritis.* Rheumatic Disease Clinics, 2019. **45**(4): p. 487-504.
333. Raza, K. and A. Filer, *The therapeutic window of opportunity in rheumatoid arthritis: does it ever close?* 2015, BMJ Publishing Group Ltd. p. 793-794.
334. Burgers, L.E., K. Raza, and A.H. Van Der Helm-Van, *Window of opportunity in rheumatoid arthritis–definitions and supporting evidence: from old to new perspectives.* RMD open, 2019. **5**(1): p. e000870.
335. Nell, V., et al., *Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis.* Rheumatology, 2004. **43**(7): p. 906-914.
336. van der Linden, M.P., et al., *Long-term impact of delay in assessment of patients with early arthritis.* Arthritis & Rheumatism, 2010. **62**(12): p. 3537-3546.
337. Lard, L.R., et al., *Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies.* The American journal of medicine, 2001. **111**(6): p. 446-451.
338. Finckh, A., et al., *Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: a meta-analysis.* Arthritis Care & Research: Official Journal of the American College of Rheumatology, 2006. **55**(6): p. 864-872.
339. Van Den Broek, M., W. Lems, and C. Allaart, *BeSt practice: the success of early-targeted treatment in rheumatoid arthritis.* Clinical and Experimental Rheumatology-Incl Supplements, 2012. **30**(4): p. S35.
340. Ten Klooster, P.M., et al., *Radiographic progression can still occur in individual patients with low or moderate disease activity in the current treat-to-target paradigm: real-world data from the Dutch Rheumatoid Arthritis Monitoring (DREAM) registry.* Arthritis research & therapy, 2019. **21**: p. 1-10.
341. Mueller, R.B., et al., *Superiority of a Treat-to-Target strategy over conventional treatment with fixed csDMARD and corticosteroids: a multi-center randomized controlled trial in RA patients with an inadequate response to conventional synthetic DMARDs, and new therapy with Certolizumab pegol.* Journal of clinical medicine, 2019. **8**(3): p. 302.
342. Verstappen, S., et al., *Intensive treatment with methotrexate in early rheumatoid arthritis: aiming for remission. Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA, an open-label*

strategy trial). Annals of the rheumatic diseases, 2007. **66**(11): p. 1443-1449.
343. Symmons, D., et al., *Patients with stable long-standing rheumatoid arthritis continue to deteriorate despite intensified treatment with traditional disease modifying anti-rheumatic drugs—results of the British Rheumatoid Outcome Study Group randomized controlled clinical trial.* Rheumatology, 2006. **45**(5): p. 558-565.
344. Harrold, L.R., et al., *Cluster-randomized trial of a behavioral intervention to incorporate a treat-to-target approach to care of US patients with rheumatoid arthritis.* Arthritis care & research, 2018. **70**(3): p. 379-387.
345. Bergsten, U., et al., *A randomized study comparing regular care with a nurse-led clinic based on tight disease activity control and person-centred care in patients with rheumatoid arthritis with moderate/high disease activity: A 6-month evaluation.* Musculoskeletal care, 2019. **17**(3): p. 215-225.
346. Fransen, J., et al., *Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial.* Annals of the rheumatic diseases, 2005. **64**(9): p. 1294-1298.
347. Van Hulst, L.T., et al., *How to improve DAS28 use in daily clinical practice?—a pilot study of a nurse-led intervention.* Rheumatology, 2010. **49**(4): p. 741-748.
348. Hock, E.S., et al., *Treat-to-target strategies in rheumatoid arthritis: a systematic review and cost-effectiveness analysis.* SN Comprehensive Clinical Medicine, 2021. **3**(3): p. 838-854.
349. Finckh, A., et al., *Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: A meta-analysis.* Arthritis Rheum, 2006. **55**(6): p. 864-72.
350. van Steenbergen, H.W., et al., *Clinical factors, anticitrullinated peptide antibodies and MRI-detected subclinical inflammation in relation to progression from clinically suspect arthralgia to arthritis.* Annals of the Rheumatic Diseases, 2016. **75**: p. 1824-1830.
351. Nam, J.L., et al., *Enriching case selection for imminent RA: the use of anti-CCP antibodies in individuals with new non-specific musculoskeletal symptoms - a cohort study.* Ann Rheum Dis, 2016. **75**(8): p. 1452-6.
352. Mankia, K., et al., *A core set of risk factors in individuals at risk of rheumatoid arthritis: a systematic literature review informing the EULAR points to consider for conducting clinical trials and observational studies in individuals at risk of rheumatoid arthritis.* RMD Open, 2021. **7**.
353. Arnett, F.C., et al., *The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 1988. **31**(3): p. 315-324.
354. Emery, P., *Pocket reference to early rheumatoid arthritis*. 2011: Springer Science & Business Media.
355. Radner, H., et al., *Performance of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: a systematic literature review.* Annals of the rheumatic diseases, 2014. **73**(1): p. 114-123.
356. Kourilovitch, M., C. Galarza-Maldonado, and E. Ortiz-Prado, *Diagnosis and classification of rheumatoid arthritis.* Journal of autoimmunity, 2014. **48**: p. 26-30.

357. Maillefert, J.-F., et al., *Prediction of response to disease modifying antirheumatic drugs in rheumatoid arthritis.* Joint Bone Spine, 2010. **77**(6): p. 558-563.
358. Brown, P.M., A.G. Pratt, and J.D. Isaacs, *Mechanism of action of methotrexate in rheumatoid arthritis, and the search for biomarkers.* Nature Reviews Rheumatology, 2016. **12**(12): p. 731-742.
359. Burska, A.N., et al., *Quantifying circulating Th17 cells by qPCR: potential as diagnostic biomarker for rheumatoid arthritis.* Rheumatology, 2019. **58**(11): p. 2015-2024.
360. López-Rodríguez, R., et al., *Replication study of polymorphisms associated with response to methotrexate in patients with rheumatoid arthritis.* Scientific reports, 2018. **8**(1): p. 7342.
361. Avdeeva, A., et al., *Regulatory T cells in patients with early untreated rheumatoid arthritis: Phenotypic changes in the course of methotrexate treatment.* Biochimie, 2020. **174**: p. 9-17.
362. Ponchel, F., et al., *Naïve T cells predict MTX induced remission in early arthritis.* Annals of the Rheumatic Diseases, 2011. **70**(Suppl 2): p. A9-A10.
363. Gupta, V., et al., *CD39 positive regulatory T cell frequency as a biomarker of treatment response to methotrexate in rheumatoid arthritis.* International Journal of Rheumatic Diseases, 2018. **21**(8): p. 1548-1556.
364. O'Dell, J.R., et al., *Validation of the methotrexate-first strategy in patients with early, poor-prognosis rheumatoid arthritis: results from a two-year randomized, double-blind trial.* Arthritis & Rheumatism, 2013. **65**(8): p. 1985-1994.
365. Möller, B., et al., *Class-switched B cells display response to therapeutic B-cell depletion in rheumatoid arthritis.* Arthritis research & therapy, 2009. **11**: p. 1-11.
366. Brezinschek, H.-P., et al., *B lymphocyte-typing for prediction of clinical response to rituximab.* Arthritis research & therapy, 2012. **14**: p. 1-9.
367. Sellam, J., et al., *Blood memory B cells are disturbed and predict the response to rituximab in patients with rheumatoid arthritis.* Arthritis & Rheumatism, 2011. **63**(12): p. 3692-3701.
368. Huang, Q.-l., et al., *Circulating biomarkers for predicting infliximab response in rheumatoid arthritis: A systematic bioinformatics analysis.* Medical Science Monitor: International Medical Journal of Experimental and Clinical Research, 2017. **23**: p. 1849.
369. Ferraccioli, G., et al., *Biomarkers of good EULAR response to the B cell depletion therapy in all seropositive rheumatoid arthritis patients: clues for the pathogenesis.* 2012.
370. Chen, D.-Y., et al., *Increasing levels of circulating Th17 cells and interleukin-17 in rheumatoid arthritis patients with an inadequate response to anti-TNF-α therapy.* Arthritis research & therapy, 2011. **13**(4): p. 1-10.
371. Hernández-Breijo, B., et al., *Remission Induced by TNF Inhibitors Plus Methotrexate is Associated with Changes in Peripheral Naïve B Cells in Patients with Rheumatoid Arthritis.* Frontiers in medicine, 2021. **8**: p. 683990.
372. Szodoray, P., et al., *Anti-citrullinated protein/peptide autoantibodies in association with genetic and environmental factors as indicators of disease outcome in rheumatoid arthritis.* Autoimmunity reviews, 2010. **9**(3): p. 140-143.

373. van der Helm-van, A.H., R.E. Toes, and T.W. Huizinga, *Genetic variants in the prediction of rheumatoid arthritis.* Annals of the rheumatic diseases, 2010. **69**(9): p. 1694-1696.
374. Balsa, A., et al., *Influence of HLA DRB1 alleles in the susceptibility of rheumatoid arthritis and the regulation of antibodies against citrullinated proteins and rheumatoid factor.* Arthritis research & therapy, 2010. **12**: p. 1-8.
375. Wysocki, T., M. Olesińska, and A. Paradowska-Gorycka, *Current understanding of an emerging role of HLA-DRB1 gene in rheumatoid arthritis–from research to clinical practice.* Cells, 2020. **9**(5): p. 1127.
376. Hirose, W., et al., *Impact of the HLA-DRB1 shared epitope on responses to treatment with tofacitinib or abatacept in patients with rheumatoid arthritis.* Arthritis Research & Therapy, 2021. **23**(1): p. 228.
377. Rigby, W., et al., *HLA-DRB1 risk alleles for RA are associated with differential clinical responsiveness to abatacept and adalimumab: data from a head-to-head, randomized, single-blind study in autoantibody-positive early RA.* Arthritis Research & Therapy, 2021. **23**: p. 1-13.
378. Murdaca, G., et al., *TNF-α gene polymorphisms: association with disease susceptibility and response to anti-TNF-α treatment in psoriatic arthritis.* Journal of Investigative Dermatology, 2014. **134**(10): p. 2503-2509.
379. Menegatti, S., E. Bianchi, and L. Rogge, *Anti-TNF therapy in spondyloarthritis and related diseases, impact on the immune system and prediction of treatment responses.* Frontiers in immunology, 2019. **10**: p. 437283.
380. Fishman, D., et al., *The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis.* The Journal of clinical investigation, 1998. **102**(7): p. 1369-1376.
381. Plant, D., A.G. Wilson, and A. Barton, *Genetic and epigenetic predictors of responsiveness to treatment in RA.* Nature Reviews Rheumatology, 2014. **10**(6): p. 329-337.
382. Nelson, M.R., et al., *The support of human genetic evidence for approved drug indications.* Nature genetics, 2015. **47**(8): p. 856-860.
383. Bek, S., et al., *Systematic review and meta-analysis: pharmacogenetics of anti-TNF treatment response in rheumatoid arthritis.* The pharmacogenomics journal, 2017. **17**(5): p. 403-411.
384. Taylor, J.C., et al., *Genome-wide association study of response to methotrexate in early rheumatoid arthritis patients.* The pharmacogenomics journal, 2018. **18**(4): p. 528-538.
385. Krintel, S.B., et al., *Prediction of treatment response to adalimumab: a double-blind placebo-controlled study of circulating microRNA in patients with early rheumatoid arthritis.* The pharmacogenomics journal, 2016. **16**(2): p. 141-146.
386. Afroz, S., et al., *A comprehensive gene expression meta-analysis identifies novel immune signatures in rheumatoid arthritis patients.* Frontiers in immunology, 2017. **8**: p. 74.
387. Rychkov, D., et al., *Cross-tissue transcriptomic analysis leveraging machine learning approaches identifies new biomarkers for rheumatoid arthritis.* Frontiers in immunology, 2021. **12**: p. 638066.

388. Talmon, M., et al., *Transcriptomic profile comparison of monocytes from rheumatoid arthritis patients in treatment with methotrexate, anti-TNFa, abatacept or tocilizumab.* PLoS One, 2023. **18**(3): p. e0282564.
389. Oliver, J., et al., *Transcriptome-wide study of TNF-inhibitor therapy in rheumatoid arthritis reveals early signature of successful treatment.* Arthritis research & therapy, 2021. **23**: p. 1-9.
390. Cheng, Q., et al., *Three hematologic/immune system-specific expressed genes are considered as the potential biomarkers for the diagnosis of early rheumatoid arthritis through bioinformatics analysis.* Journal of translational medicine, 2021. **19**: p. 1-15.
391. Toro-Domínguez, D., P. Carmona-Sáez, and M.E. Alarcón-Riquelme, *Shared signatures between rheumatoid arthritis, systemic lupus erythematosus and Sjögren's syndrome uncovered through gene expression meta-analysis.* Arthritis research & therapy, 2014. **16**: p. 1-8.
392. Tyagi, N., K. Mehla, and D. Gupta, *Deciphering novel common gene signatures for rheumatoid arthritis and systemic lupus erythematosus by integrative analysis of transcriptomic profiles.* Plos one, 2023. **18**(3): p. e0281637.
393. Yasumizu, Y., et al., *Single-cell transcriptome landscape of circulating CD4+ T cell populations in autoimmune diseases.* Cell Genomics, 2024.
394. Lien, H.J.T., et al., *Single-cell resolution of longitudinal blood transcriptome profiles in rheumatoid arthritis, systemic lupus erythematosus and healthy control pregnancies.* Annals of the Rheumatic Diseases, 2024. **83**(3): p. 300-311.
395. Qin, H., T. Niu, and J. Zhao, *Identifying multi-omics causers and causal pathways for complex traits.* Frontiers in genetics, 2019. **10**: p. 386438.
396. Kiernan, U.A., *Biomarker rediscovery in diagnostics.* Expert Opinion on Medical Diagnostics, 2008. **2**(12): p. 1391-1400.
397. Shruthi, B.S. and P. Vinodhkumar, *Proteomics: A new perspective for cancer.* Advanced biomedical research, 2016. **5**(1): p. 67.
398. Vuong, H., et al., *Functional consequences of somatic mutations in cancer using protein pocket-based prioritization approach.* Genome medicine, 2014. **6**: p. 1-14.
399. Engin, H.B., J.F. Kreisberg, and H. Carter, *Structure-based analysis reveals cancer missense mutations target protein interaction interfaces.* PloS one, 2016. **11**(4): p. e0152929.
400. McArdle, A., et al., *Developing clinically relevant biomarkers in inflammatory arthritis: a multiplatform approach for serum candidate protein discovery.* PROTEOMICS–Clinical Applications, 2016. **10**(6): p. 691-698.
401. Dayon, L., O. Cominetti, and M. Affolter, *Proteomics of human biological fluids for biomarker discoveries: technical advances and recent applications.* Expert review of proteomics, 2022. **19**(2): p. 131-151.
402. Sobolev, V.V., et al., *Proteomic studies of psoriasis.* Biomedicines, 2022. **10**(3): p. 619.
403. Brown, M.A., Z. Li, and K.-A.L. Cao, *Biomarker development for axial spondyloarthritis.* Nature Reviews Rheumatology, 2020. **16**(8): p. 448-463.
404. O'Neil, L.J., et al., *Proteomic approaches to defining remission and the risk of relapse in rheumatoid arthritis.* Frontiers in Immunology, 2021. **12**: p. 729681.

405. Mallick, P. and B. Kuster, *Proteomics: a pragmatic perspective.* Nature biotechnology, 2010. **28**(7): p. 695-709.
406. Chang, X., et al., *Identification of proteins with increased expression in rheumatoid arthritis synovial tissues.* The Journal of rheumatology, 2009. **36**(5): p. 872-880.
407. Ardle, A.M., et al., *Identification and Evaluation of Serum Protein Biomarkers Which Differentiate Psoriatic from Rheumatoid Arthritis.* medRxiv, 2020: p. 2020.06. 23.20138552.
408. Tasaki, S., et al., *Multi-omics monitoring of drug response in rheumatoid arthritis in pursuit of molecular remission.* Nature communications, 2018. **9**(1): p. 2755.
409. Chen, S.-F., et al., *Tailored therapeutic decision of rheumatoid arthritis using proteomic strategies: how to start and when to stop?* Clinical Proteomics, 2023. **20**(1): p. 22.
410. Ling, S.F., et al., *A proteomics study of rheumatoid arthritis patients on etanercept identifies putative biomarkers associated with clinical outcome measures.* Rheumatology, 2024. **63**(4): p. 1015-1021.
411. O'Neil, L.J., et al., *Association of a serum protein signature with rheumatoid arthritis development.* Arthritis & Rheumatology, 2021. **73**(1): p. 78-88.
412. Ortea, I., et al., *Discovery of serum proteomic biomarkers for prediction of response to infliximab (a monoclonal anti-TNF antibody) treatment in rheumatoid arthritis: an exploratory analysis.* Journal of proteomics, 2012. **77**: p. 372-382.
413. Gardette, A., et al., *High anti-CCP antibody titres predict good response to rituximab in patients with active rheumatoid arthritis.* Joint Bone Spine, 2014. **81**(5): p. 416-420.
414. Gottenberg, J., et al., *Positivity for anti-cyclic citrullinated peptide is associated with a better response to abatacept: data from the 'Orencia and Rheumatoid Arthritis' registry.* Annals of the rheumatic diseases, 2012. **71**(11): p. 1815-1819.
415. Sokolova, M.V., et al., *Antibodies against citrullinated proteins of IgA isotype are associated with progression to rheumatoid arthritis in individuals at-risk.* RMD open, 2023. **9**(1): p. e002705.
416. Cunningham, K.Y., et al., *Patients with ACPA-positive and ACPA-negative rheumatoid arthritis show different serological autoantibody repertoires and autoantibody associations with disease activity.* Scientific reports, 2023. **13**(1): p. 5360.
417. Chandran, V., et al., *Soluble biomarkers differentiate patients with psoriatic arthritis from those with psoriasis without arthritis.* Rheumatology, 2010. **49**(7): p. 1399-1405.
418. Schett, G., et al., *Oral apremilast in the treatment of active psoriatic arthritis: results of a multicenter, randomized, double-blind, placebo-controlled study.* Arthritis & Rheumatism, 2012. **64**(10): p. 3156-3167.
419. Mease, P.J., et al., *Brodalumab, an anti-IL17RA monoclonal antibody, in psoriatic arthritis.* New England Journal of Medicine, 2014. **370**(24): p. 2295-2306.
420. Kavanaugh, A., et al., *Ustekinumab, an anti-IL-12/23 p40 monoclonal antibody, inhibits radiographic progression in patients with active psoriatic arthritis: results of an integrated analysis of radiographic data from the phase 3, multicentre, randomised, double-blind, placebo-controlled*

PSUMMIT-1 and PSUMMIT-2 trials. Annals of the rheumatic diseases, 2014. **73**(6): p. 1000-1006.
421. Liu, S., et al., *Discovery of potential serum protein biomarkers in ankylosing spondylitis using tandem mass tag-based quantitative proteomics.* Journal of proteome research, 2020. **19**(2): p. 864-872.
422. Luan, H., et al., *Serum metabolomic and lipidomic profiling identifies diagnostic biomarkers for seropositive and seronegative rheumatoid arthritis patients.* Journal of translational medicine, 2021. **19**: p. 1-10.
423. Koh, J.H., et al., *Lipidome profile predictive of disease evolution and activity in rheumatoid arthritis.* Experimental & Molecular Medicine, 2022. **54**(2): p. 143-155.
424. Song, X. and Q. Lin, *Genomics, transcriptomics and proteomics to elucidate the pathogenesis of rheumatoid arthritis.* Rheumatology international, 2017. **37**: p. 1257-1265.
425. Moosavi, A. and A.M. Ardekani, *Role of epigenetics in biology and human diseases.* Iranian biomedical journal, 2016. **20**(5): p. 246.
426. Farsetti, A., B. Illi, and C. Gaetano, *How epigenetics impacts on human diseases.* European Journal of Internal Medicine, 2023.
427. Ponchel, F. and R.P. , and Thomas Sargent, *Differential CpG DNA Methylation: qMSP a Promising New Approach toward Clinical Utility as Biomarkers* Annals of Clinical and Laboratory Research, 2023. **11**(S4): p. 1-4.
428. Pitaksalee, R., et al., *Differential CpG DNA methylation in peripheral naïve CD4+ T-cells in early rheumatoid arthritis patients.* Clinical epigenetics, 2020. **12**(1): p. 1-16.
429. Nile, C.J., et al., *Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis.* Arthritis & Rheumatism, 2008. **58**(9): p. 2686-2693.
430. Ishida, K., et al., *Interleukin-6 gene promoter methylation in rheumatoid arthritis and chronic periodontitis.* Journal of periodontology, 2012. **83**(7): p. 917-925.
431. Nakano, K., et al., *DNA methylome signature in rheumatoid arthritis.* Annals of the rheumatic diseases, 2013. **72**(1): p. 110-117.
432. Fu, L.-h., et al., *Hypomethylation of proximal CpG motif of interleukin-10 promoter regulates its expression in human rheumatoid arthritis.* Acta Pharmacologica Sinica, 2011. **32**(11): p. 1373-1380.
433. Liu, C.-C., et al., *Global DNA methylation, DNMT1, and MBD2 in patients with rheumatoid arthritis.* Immunology letters, 2011. **135**(1-2): p. 96-99.
434. Plant, D., et al., *Differential methylation as a biomarker of response to etanercept in patients with rheumatoid arthritis.* Arthritis & Rheumatology, 2016. **68**(6): p. 1353-1360.
435. Karouzakis, E., et al., *DNA methylation regulates the expression of CXCL12 in rheumatoid arthritis synovial fibroblasts.* Genes & Immunity, 2011. **12**(8): p. 643-652.
436. Rodríguez-Ubreva, J., et al., *Inflammatory cytokines shape a changing DNA methylome in monocytes mirroring disease activity in rheumatoid arthritis.* Annals of the Rheumatic Diseases, 2019. **78**(11): p. 1505-1516.
437. Lamot, L., et al. *Hypermethylation Of NLRP3 Promoter Region Could Be Responsible For Decreased Gene Expression, Inflammasome Malfunction And Gut Dysbiosis In Early Phase Juvenile Spondyloarthritis*. in *Nepoznat skup*. 2018.

438. Lamot, L., et al., *Epigenetic alterations in juvenile spondyloarthritis patients: a preliminary study of selected genes promoter methylation and silencing.* SN comprehensive clinical medicine, 2019. **1**: p. 496-501.
439. de Andres, M.C., et al., *Assessment of global DNA methylation in peripheral blood cell subpopulations of early rheumatoid arthritis before and after methotrexate.* Arthritis research & therapy, 2015. **17**: p. 1-9.
440. Wilson, A.S., B.E. Power, and P.L. Molloy, *DNA hypomethylation and human diseases.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2007. **1775**(1): p. 138-162.
441. Davidović, R.S., et al., *Methylation-specific PCR: four steps in primer design.* Central European journal of biology, 2014. **9**: p. 1127-1139.
442. Kim, Y.-I., et al., *DNA hypomethylation in inflammatory arthritis: reversal with methotrexate.* Journal of Laboratory and Clinical Medicine, 1996. **128**(2): p. 165-172.
443. Huber, L.C., et al., *Histone deacetylase/acetylase activity in total synovial tissue derived from rheumatoid arthritis and osteoarthritis patients.* Arthritis & Rheumatism, 2007. **56**(4): p. 1087-1093.
444. Grabiec, A.M., et al., *Histone deacetylase inhibitors suppress inflammatory activation of rheumatoid arthritis patient synovial macrophages and tissue.* The journal of immunology, 2010. **184**(5): p. 2718-2728.
445. Miao, C.-g., et al., *New advances of DNA methylation and histone modifications in rheumatoid arthritis, with special emphasis on MeCP2.* Cellular signalling, 2013. **25**(4): p. 875-882.
446. Lyko, F., *The DNA methyltransferase family: a versatile toolkit for epigenetic regulation.* Nature Reviews Genetics, 2018. **19**(2): p. 81-92.
447. Nakano, K., D.L. Boyle, and G.S. Firestein, *Regulation of DNA methylation in rheumatoid arthritis synoviocytes.* The journal of immunology, 2013. **190**(3): p. 1297-1303.
448. Wada, T.T., et al., *Aberrant histone acetylation contributes to elevated interleukin-6 production in rheumatoid arthritis synovial fibroblasts.* Biochemical and biophysical research communications, 2014. **444**(4): p. 682-686.
449. Grabiec, A.M., et al., *Histone deacetylase inhibitors suppress rheumatoid arthritis fibroblast-like synoviocyte and macrophage IL-6 production by accelerating mRNA decay.* Annals of the rheumatic diseases, 2012. **71**(3): p. 424-431.
450. Wohlers, I., L. Bertram, and C.M. Lill, *Evidence for a potential role of miR-1908-5p and miR-3614-5p in autoimmune disease risk using integrative bioinformatics.* Journal of Autoimmunity, 2018. **94**: p. 83-89.
451. Stanczyk, J., et al., *Altered expression of microRNA-203 in rheumatoid arthritis synovial fibroblasts and its role in fibroblast activation.* Arthritis & Rheumatism, 2011. **63**(2): p. 373-381.
452. Ciechomska, M., et al., *Changes in MiRNA-5196 expression as a potential biomarker of anti-TNF-α therapy in rheumatoid arthritis and ankylosing spondylitis patients.* Archivum immunologiae et therapiae experimentalis, 2018. **66**: p. 389-397.
453. Evangelatos, G., et al., *MicroRNAs in rheumatoid arthritis: From pathogenesis to clinical impact.* Autoimmunity reviews, 2019. **18**(11): p. 102391.

454. Walther, K. and L.N. Schulte, *The role of lncRNAs in innate immunity and inflammation.* RNA biology, 2021. **18**(5): p. 587-603.
455. Huang, W., et al., *LncRNAs and rheumatoid arthritis: from identifying mechanisms to clinical investigation.* Frontiers in Immunology, 2022. **12**: p. 807738.
456. Niimoto, T., et al., *MicroRNA-146a expresses in interleukin-17 producing T cells in rheumatoid arthritis patients.* BMC musculoskeletal disorders, 2010. **11**: p. 1-11.
457. Li, J., et al., *Altered microRNA expression profile with miR-146a upregulation in CD4+ T cells from patients with rheumatoid arthritis.* Arthritis research & therapy, 2010. **12**: p. 1-12.
458. Churov, A.V., E.K. Oleinik, and M. Knip, *MicroRNAs in rheumatoid arthritis: altered expression and diagnostic potential.* Autoimmunity reviews, 2015. **14**(11): p. 1029-1037.
459. Duroux-Richard, I., C. Jorgensen, and F. Apparailly, *What do microRNAs mean for rheumatoid arthritis?* Arthritis & Rheumatism, 2012. **64**(1): p. 11-20.
460. Li, X., F. Tian, and F. Wang, *Rheumatoid arthritis-associated microRNA-155 targets SOCS1 and upregulates TNF-α and IL-1β in PBMCs.* International journal of molecular sciences, 2013. **14**(12): p. 23910-23921.
461. Tang, J., S. Yi, and Y. Liu, *Long non-coding RNA PVT1 can regulate the proliferation and inflammatory responses of rheumatoid arthritis fibroblast-like synoviocytes by targeting microRNA-145-5p.* Human Cell, 2020. **33**(4): p. 1081-1090.
462. Bi, X., et al., *LncRNA PICSAR promotes cell proliferation, migration and invasion of fibroblast-like synoviocytes by sponging miRNA-4701-5p in rheumatoid arthritis.* EBioMedicine, 2019. **50**: p. 408-420.
463. Scrivo, R., et al., *The immunology of rheumatoid arthritis.* Annals of the New York Academy of Sciences, 2007. **1108**(1): p. 312-322.
464. Lübbers, J., et al., *Changes in peripheral blood lymphocyte subsets during arthritis development in arthralgia patients.* Arthritis research & therapy, 2016. **18**(1): p. 1-10.
465. Chalan, P., et al., *Circulating CD4+ CD161+ T lymphocytes are increased in seropositive arthralgia patients but decreased in patients with newly diagnosed rheumatoid arthritis.* PloS one, 2013. **8**(11): p. e79370.
466. Chalan, P., et al., *Altered natural killer cell subsets in seropositive arthralgia and early rheumatoid arthritis are associated with autoantibody status.* The Journal of rheumatology, 2016. **43**(6): p. 1008-1016.
467. Jonsson, A.H., et al., *Granzyme K+ CD8 T cells form a core population in inflamed human tissue.* Science translational medicine, 2022. **14**(649): p. eabo0686.
468. Wherry, E.J., et al., *Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment.* Journal of virology, 2003. **77**(8): p. 4911-4927.
469. Kurachi, M., et al., *Chemokine receptor CXCR3 facilitates CD8+ T cell differentiation into short-lived effector cells leading to memory degeneration.* Journal of experimental medicine, 2011. **208**(8): p. 1605-1620.
470. Kurts, C., et al., *Class I–restricted cross-presentation of exogenous self-antigens leads to deletion of autoreactive CD8+ T cells.* The Journal of experimental medicine, 1997. **186**(2): p. 239-245.

471. Sundaresan, B., et al., *The role of viral infections in the onset of autoimmune diseases.* Viruses, 2023. **15**(3): p. 782.
472. Balandraud, N. and J. Roudier, *Epstein-Barr virus and rheumatoid arthritis.* Joint Bone Spine, 2018. **85**(2): p. 165-170.
473. Joo, Y.B., et al., *Respiratory viral infections and the risk of rheumatoid arthritis.* Arthritis research & therapy, 2019. **21**(1): p. 1-8.
474. Zhao, J., et al., *Apoptosis, autophagy, NETosis, necroptosis, and pyroptosis mediated programmed cell death as targets for innovative therapy in rheumatoid arthritis.* Frontiers in Immunology, 2021. **12**: p. 809806.
475. Nanki, T., et al., *Chemokine receptor expression and functional effects of chemokines on B cells: implication in the pathogenesis of rheumatoid arthritis.* Arthritis research & therapy, 2009. **11**: p. 1-11.
476. Van Baarsen, L., et al., *The cellular composition of lymph nodes in the earliest phase of inflammatory arthritis.* Annals of the rheumatic diseases, 2013. **72**(8): p. 1420-1424.
477. Janssen, K.M., et al., *Regulatory CD4+ T-cell subsets and anti-citrullinated protein antibody repertoire: potential biomarkers for arthritis development in seropositive arthralgia patients?* PLoS One, 2016. **11**(9): p. e0162101.
478. Seddiki, N., et al., *Expression of interleukin (IL)-2 and IL-7 receptors discriminates between human regulatory and activated T cells.* The Journal of experimental medicine, 2006. **203**(7): p. 1693-1700.
479. Liu, W., et al., *CD127 expression inversely correlates with FoxP3 and suppressive function of human CD4+ T reg cells.* The Journal of experimental medicine, 2006. **203**(7): p. 1701-1711.
480. Yu, N., et al., *CD4+ CD25+ CD127 low/− T cells: a more specific Treg population in human peripheral blood.* Inflammation, 2012. **35**: p. 1773-1780.
481. Ward-Hartstonge, K.A. and R.A. Kemp, *Regulatory T-cell heterogeneity and the cancer immune response.* Clinical & translational immunology, 2017. **6**(9): p. e154.
482. Cascão, R., et al., *Identification of a cytokine network sustaining neutrophil and Th17 activation in untreated early rheumatoid arthritis.* Arthritis research & therapy, 2010. **12**(5): p. 1-8.
483. Kotake, S., et al., *Elevated ratio of Th17 cell-derived Th1 cells (CD161+ Th1 Cells) to CD161+ Th17 cells in peripheral blood of early-onset rheumatoid arthritis patients.* BioMed research international, 2016. **2016**.
484. Daïen, C.I., et al., *High levels of natural killer cells are associated with response to tocilizumab in patients with severe rheumatoid arthritis.* Rheumatology, 2015. **54**(4): p. 601-608.
485. Lawson, C.A., et al., *Early rheumatoid arthritis is associated with a deficit in the CD4+CD25high regulatory T cell population in peripheral blood.* Rheumatology, 2006. **45**(10): p. 1210-7.
486. Maddur, M.S., et al., *Th17 cells: biology, pathogenesis of autoimmune and inflammatory diseases, and therapeutic strategies.* The American journal of pathology, 2012. **181**(1): p. 8-18.
487. Paulissen, S.M., et al., *CCR6+ Th cell populations distinguish ACPA positive from ACPA negative rheumatoid arthritis.* Arthritis research & therapy, 2015. **17**(1): p. 1-9.

488. Kahmini, F.R., et al., *Emerging therapeutic potential of regulatory T (Treg) cells for rheumatoid arthritis: new insights and challenges.* International Immunopharmacology, 2022. **108**: p. 108858.
489. Cribbs, A.P., et al., *Treg cell function in rheumatoid arthritis is compromised by CTLA-4 promoter methylation resulting in a failure to activate the indoleamine 2, 3-dioxygenase pathway.* Arthritis & rheumatology, 2014. **66**(9): p. 2344-2354.
490. Kawashiri, S.-Y., et al., *CD4+ CD25highCD127low/-Treg cell frequency from peripheral blood correlates with disease activity in patients with rheumatoid arthritis.* The Journal of rheumatology, 2011. **38**(12): p. 2517-2521.
491. Saleem, B., et al., *Patients with RA in remission on TNF blockers: when and in whom can TNF blocker therapy be stopped?* Annals of the rheumatic diseases, 2010. **69**(9): p. 1636-1642.
492. Han, G.M., et al., *CD4+ CD25high T cell numbers are enriched in the peripheral blood of patients with rheumatoid arthritis.* Cellular immunology, 2008. **253**(1-2): p. 92-101.
493. Paradowska-Gorycka, A., et al., *Th17/Treg-related transcriptional factor expression and cytokine profile in patients with rheumatoid arthritis.* Frontiers in Immunology, 2020. **11**: p. 572858.
494. Yang, M., et al., *Helios but not CD226, TIGIT and Foxp3 is a potential marker for CD4+ Treg cells in patients with rheumatoid arthritis.* Cellular physiology and biochemistry: international journal of experimental cellular physiology, biochemistry, and pharmacology, 2019. **52**(5): p. 1178.
495. McGovern, J.L., et al., *Th17 cells are restrained by Treg cells via the inhibition of interleukin-6 in patients with rheumatoid arthritis responding to anti–tumor necrosis factor antibody therapy.* Arthritis & Rheumatism, 2012. **64**(10): p. 3129-3138.
496. Niu, H. and L. Hao, *AB0237 CHANGES AND CLINICAL SIGNIFICANCES OF PERIPHERAL REGULATORY T CELLS IN RHEUMATOID ARTHRITIS PATIENTS WITH CARDIOVASCULAR DISEASE*. 2022, BMJ Publishing Group Ltd.
497. Yang, M., et al., *Regulatory B cells in autoimmune diseases.* Cellular & molecular immunology, 2013. **10**(2): p. 122-132.
498. Peng, B., Y. Ming, and C. Yang, *Regulatory B cells: the cutting edge of immune tolerance in kidney transplantation.* Cell death & disease, 2018. **9**(2): p. 109.
499. Sakkas, L.I., *Regulatory B cells in autoimmune rheumatic diseases.* Mediterranean Journal of Rheumatology, 2017. **28**(2): p. 75-79.
500. Mauri, C. and M. Menon, *Human regulatory B cells in health and disease: therapeutic potential.* The Journal of clinical investigation, 2017. **127**(3): p. 772-779.
501. Cherukuri, A., et al., *Immunologic human renal allograft injury associates with an altered IL-10/TNF-α expression ratio in regulatory B cells.* Journal of the American Society of Nephrology: JASN, 2014. **25**(7): p. 1575.
502. Simon, Q. and S. Hillion, *Regulatory B Cells.* Translational Immunology: Mechanisms and Pharmacologic Approaches, 2015: p. 247.
503. Tudhope, S.J., et al., *Profound invariant natural killer T-cell deficiency in inflammatory arthritis.* Annals of the rheumatic diseases, 2010. **69**(10): p. 1873-1879.

504. Souto-Carneiro, M.M., et al., *Alterations in peripheral blood memory B cells in patients with active rheumatoid arthritis are dependent on the action of tumour necrosis factor.* Arthritis research & therapy, 2009. **11**(3): p. 1-12.
505. Ota, Y., et al., *Generation mechanism of RANKL+ effector memory B cells: relevance to the pathogenesis of rheumatoid arthritis.* Arthritis research & therapy, 2016. **18**(1): p. 1-10.
506. Burmester, G.R. and J.E. Pope, *Novel treatment strategies in rheumatoid arthritis.* The Lancet, 2017. **389**(10086): p. 2338-2348.
507. Buckley, C., *EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update.* Annals of the Rheumatic Diseases, 2017.
508. McInnes, I.B. and J.R. O'Dell, *State-of-the-art: rheumatoid arthritis.* Annals of the rheumatic diseases, 2010. **69**(11): p. 1898-1906.
509. Smolen, J.S., et al., *EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update.* Annals of the rheumatic diseases, 2014. **73**(3): p. 492-509.
510. Weinblatt, M., *Efficacy of methotrexate in rheumatoid arthritis.* Rheumatology, 1995. **34**(suppl_2): p. 43-48.
511. Katchamart, W., et al., *Efficacy and toxicity of methotrexate (MTX) monotherapy versus MTX combination therapy with non-biological disease-modifying antirheumatic drugs in rheumatoid arthritis: a systematic review and meta-analysis.* Annals of the rheumatic diseases, 2009. **68**(7): p. 1105-1112.
512. Peres, R.S., et al., *Low expression of CD39 on regulatory T cells as a biomarker for resistance to methotrexate therapy in rheumatoid arthritis.* Proceedings of the National Academy of Sciences, 2015. **112**(8): p. 2509-2514.
513. Breedveld, F.C., et al., *The PREMIER study: a multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2006. **54**(1): p. 26-37.
514. Emery, P., et al., *Comparison of methotrexate monotherapy with a combination of methotrexate and etanercept in active, early, moderate to severe rheumatoid arthritis (COMET): a randomised, double-blind, parallel treatment trial.* The Lancet, 2008. **372**(9636): p. 375-382.
515. Prasad, S., et al., *Multidrug resistance protein-1 expression, function and polymorphisms in patients with rheumatoid arthritis not responding to methotrexate.* International Journal of Rheumatic Diseases, 2014. **17**(8): p. 878-886.
516. Hider, S.L., et al., *Can clinical factors at presentation be used to predict outcome of treatment with methotrexate in patients with early inflammatory polyarthritis?* Annals of the rheumatic diseases, 2009. **68**(1): p. 57-62.
517. Scirè, C.A., et al., *Reduction of long-term disability in inflammatory polyarthritis by early and persistent suppression of joint inflammation: results from the Norfolk Arthritis Register.* Arthritis care & research, 2011. **63**(7): p. 945-952.

518. Kanjana, K., et al., *Inhibitory activity of FOXP3+ regulatory T cells reveals high specificity for displaying immune tolerance in remission state rheumatoid arthritis.* Scientific reports, 2020. **10**(1): p. 19789.
519. Brown, P., et al., *OP0033 REGULATORY T CELL CD39 EXPRESSION AS A PREDICTOR OF EARLY REMISSION-INDUCTION WITH METHOTREXATE IN NEW-ONSET RHEUMATOID ARTHRITIS*. 2021, BMJ Publishing Group Ltd.
520. Ponchel, F., et al., *An immunological biomarker to predict MTX response in early RA.* Annals of the rheumatic diseases, 2014. **73**(11): p. 2047-2053.
521. UoL. *Targeted Treatment Early With Etanercept + Methotrexate vs.T2T Care for DMARD-naïve Early RA Patients Based on naïve T-cell Stratification (TEEMS)*. 2019 20 November 2022]; Available from: https://classic.clinicaltrials.gov/ct2/show/NCT03813771.
522. Fedele, A.L., et al., *Memory B cell subsets and plasmablasts are lower in early than in long-standing rheumatoid arthritis.* Bmc Immunology, 2014. **15**: p. 1-9.
523. Fortea-Gordo, P., et al., *Circulating CD19+ CD24hiCD38hi regulatory B cells as biomarkers of response to methotrexate in early rheumatoid arthritis.* Rheumatology, 2020. **59**(10): p. 3081-3091.
524. Chara, L., et al., *The number of circulating monocytes as biomarkers of the clinical response to methotrexate in untreated patients with rheumatoid arthritis.* Journal of translational medicine, 2015. **13**(1): p. 1-10.
525. Cooper, D.L., et al., *FcγRIIIa expression on monocytes in rheumatoid arthritis: role in immune-complex stimulated TNF production and non-response to methotrexate therapy.* PloS one, 2012. **7**(1): p. e28918.
526. Aldridge, J., et al., *Blood PD-1+ TFh and CTLA-4+ CD4+ T cells predict remission after CTLA-4Ig treatment in early rheumatoid arthritis.* Rheumatology, 2022. **61**(3): p. 1233-1242.
527. F. Tastekin, M.L., P. Ravi, H. Ng,, M. Buch, P. emery, f. ponchel, *Value of T-cells subsets in predicitign responses to MTX and anti-TNF in early RA*, in *43rd European Workshop for Rheumatology Research* 2024, EWRR Genoa, Italy.
528. Salomon, S., et al., *Th17 and CD24hiCD27+ regulatory B lymphocytes are biomarkers of response to biologics in rheumatoid arthritis.* Arthritis research & therapy, 2017. **19**(1): p. 1-10.
529. Sałkowska, A., et al., *Identification of novel molecular markers of human Th17 cells.* Cells, 2020. **9**(7): p. 1611.
530. Aggarwal, S., et al., *Interleukin-23 promotes a distinct CD4 T cell activation state characterized by the production of interleukin-17.* Journal of Biological Chemistry, 2003. **278**(3): p. 1910-1914.
531. Bedoya, S.K., et al., *Isolation and th17 differentiation of naive CD4 T lymphocytes.* JoVE (Journal of Visualized Experiments), 2013(79): p. e50765.
532. De Wit, J., et al., *CD5 costimulation induces stable Th17 development by promoting IL-23R expression and sustained STAT3 activation.* Blood, The Journal of the American Society of Hematology, 2011. **118**(23): p. 6107-6114.
533. Valmori, D., et al., *A peripheral circulating compartment of natural naive CD4+ Tregs.* The Journal of clinical investigation, 2005. **115**(7): p. 1953-1962.

534. Citro, A., et al., *CD8+ T cells specific to apoptosis-associated antigens predict the response to tumor necrosis factor inhibitor therapy in rheumatoid arthritis.* PloS one, 2015. **10**(6): p. e0128607.
535. Daien, C.I., et al., *High levels of memory B cells are associated with response to a first tumor necrosis factor inhibitor in patients with rheumatoid arthritis in a longitudinal prospective study.* Arthritis research & therapy, 2014. **16**(2): p. 1-10.
536. Laura, M., et al., *Low Percentage of Signal Regulatory Protein α/β+ Memory B Cells in Blood Predicts Development of Anti-drug Antibodies (ADA) in Adalimumab-Treated Rheumatoid Arthritis Patients.* Frontiers in immunology, 2018. **9**.
537. Solomon, G., *Immunogenicity: implications for rheumatoid arthritis treatment.* Bulletin of the NYU Hospital for Joint Diseases, 2013. **71**(3): p. 200.
538. Takeuchi, T., *Biomarkers as a treatment guide in rheumatoid arthritis.* Clinical Immunology, 2018. **186**: p. 59-62.
539. Bartelds, G.M., et al., *Development of antidrug antibodies against adalimumab and association with disease activity and treatment failure during long-term follow-up.* Jama, 2011. **305**(14): p. 1460-1468.
540. Shankar, G., et al., *Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides—harmonized terminology and tactical recommendations.* The AAPS journal, 2014. **16**: p. 658-673.
541. Karmiris, K., et al., *Influence of trough serum levels and immunogenicity on long-term outcome of adalimumab therapy in Crohn's disease.* Gastroenterology, 2009. **137**(5): p. 1628-1640.
542. Calabresi, P., et al., *The incidence and significance of anti-natalizumab antibodies: results from AFFIRM and SENTINEL.* Neurology, 2007. **69**(14): p. 1391-1403.
543. West, R., et al., *Immunogenicity negatively influences the outcome of adalimumab treatment in Crohn's disease.* Alimentary pharmacology & therapeutics, 2008. **28**(9): p. 1122-1126.
544. Radstake, T.R., et al., *Formation of antibodies against infliximab and adalimumab strongly correlates with functional drug levels and clinical responses in rheumatoid arthritis.* Annals of the rheumatic diseases, 2009. **68**(11): p. 1739-1745.
545. Bendtzen, K., et al., *Individualized monitoring of drug bioavailability and immunogenicity in rheumatoid arthritis patients treated with the tumor necrosis factor α inhibitor infliximab.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2006. **54**(12): p. 3782-3789.
546. Rodríguez-Martín, E., et al., *Blood lymphocyte subsets for early identification of non-remission to TNF inhibitors in rheumatoid arthritis.* Frontiers in immunology, 2020. **11**: p. 1913.
547. Chara, L., et al., *Monocyte populations as markers of response to adalimumab plus MTX in rheumatoid arthritis.* Arthritis research & therapy, 2012. **14**(4): p. 1-11.
548. Daien, C., et al., *THU0187 Memory B Cells Produce More TNF Than Other B Cell Subtypes and Their Percentage at Baseline Could Help to Predict Response to TNF-Inhibitors.* Annals of the Rheumatic Diseases, 2013. **72**(Suppl 3): p. A226-A227.

549. Kremer, J.M., et al., *Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig.* New England Journal of Medicine, 2003. **349**(20): p. 1907-1915.
550. Emery, P., *The therapeutic potential of costimulatory blockade with CTLA4Ig in rheumatoid arthritis.* Expert opinion on investigational drugs, 2003. **12**(4): p. 673-681.
551. Emery, P., et al., *Impact of T-cell costimulation modulation in patients with undifferentiated inflammatory arthritis or very early rheumatoid arthritis: a clinical and imaging study of abatacept (the ADJUST trial).* Annals of the rheumatic diseases, 2010. **69**(3): p. 510-516.
552. Inamo, J., et al., *High serum IgA and activated Th17 and Treg predict the efficacy of abatacept in patients with early, seropositive rheumatoid arthritis.* Clinical Rheumatology, 2021. **40**: p. 3615-3626.
553. Scarsi, M. and T. Ziglioli, *Baseline numbers of circulating CD28-negative T cells may predict clinical response to abatacept in patients with rheumatoid arthritis.* The Journal of Rheumatology, 2011. **38**(10): p. 2105-2111.
554. Gazeau, P., et al., *Memory B cells and response to abatacept in rheumatoid arthritis.* Clinical reviews in allergy & immunology, 2017. **53**(2): p. 166-176.
555. 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 & Rheumatism: Official Journal of the American College of Rheumatology, 2006. **54**(5): p. 1390-1400.
556. Chatzidionysiou, K., et al., *Highest clinical effectiveness of rituximab in autoantibody-positive patients with rheumatoid arthritis and in those for whom no more than one previous TNF antagonist has failed: pooled data from 10 European registries.* Annals of the rheumatic diseases, 2011. **70**(9): p. 1575-1580.
557. Rovin, B.H., et al., *Efficacy and safety of rituximab in patients with active proliferative lupus nephritis: the Lupus Nephritis Assessment with Rituximab study.* Arthritis & Rheumatism, 2012. **64**(4): p. 1215-1226.
558. Merrill, J.T., et al., *Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2010. **62**(1): p. 222-233.
559. McInnes, I.B. and G. Schett, *Pathogenetic insights from the treatment of rheumatoid arthritis.* The Lancet, 2017. **389**(10086): p. 2328-2337.
560. Stradner, M.H., et al., *A combination of cellular biomarkers predicts failure to respond to rituximab in rheumatoid arthritis: a 24-week observational study.* Arthritis research & therapy, 2016. **18**(1): p. 1-8.
561. Vital, E.M., et al., *Reduced-dose rituximab in rheumatoid arthritis: efficacy depends on degree of B cell depletion.* Arthritis & Rheumatism, 2011. **63**(3): p. 603-608.
562. Nakou, M., et al., *Rituximab therapy reduces activated B cells in both the peripheral blood and bone marrow of patients with rheumatoid arthritis: depletion of memory B cells correlates with clinical response.* Arthritis research & therapy, 2009. **11**(4): p. 1-8.

563. Adlowitz, D.G., 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): p. e0128269.
564. Tony, H., et al., *Combination of B cell biomarkers as independent predictors of response in patients with RA treated with rituximab.* Clin Exp Rheumatol, 2015. **33**(6): p. 887-894.
565. Parietti, V., et al., *Rituximab treatment overcomes reduction of regulatory iNKT cells in patients with rheumatoid arthritis.* Clinical Immunology, 2010. **134**(3): p. 331-339.
566. Lurati, A., et al., *NK cell count as predictor of clinical response in patients with rheumatoid arthritis treated with rituximab.* Biologics: Targets and Therapy, 2012: p. 83-87.
567. Vital, E., et al., *Management of nonresponse to rituximab in rheumatoid arthritis: predictors and outcome of re-treatment.* Arthritis & Rheumatism, 2010. **62**(5): p. 1273-1279.
568. Dass, S., et al., *Highly sensitive B cell analysis predicts response to rituximab therapy in rheumatoid arthritis.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2008. **58**(10): p. 2993-2999.
569. Leandro, M., et al., *Repopulation of B lymphocytes in peripheral blood following B lymphocyte depletion with rituximab in rheumatoid arthritis.* Arthritis Rheum, 2006. **54**: p. 613-620.
570. Emery, P., et al., *IL-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial.* Annals of the rheumatic diseases, 2008. **67**(11): p. 1516-1523.
571. Emery, P., et al., *IL-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial.(vol 67, pg 1516, 2008).* Annals of the rheumatic diseases, 2009. **68**(2): p. 296.
572. Weinblatt, M.E., et al., *A randomized phase II b study of mavrilimumab and golimumab in rheumatoid arthritis.* Arthritis & rheumatology, 2018. **70**(1): p. 49-59.
573. Atzeni, F., et al., *Long-term anti-TNF therapy and the risk of serious infections in a cohort of patients with rheumatoid arthritis: comparison of adalimumab, etanercept and infliximab in the GISEA registry.* Autoimmunity reviews, 2012. **12**(2): p. 225-229.
574. Tak, P.P., *A personalized medicine approach to biologic treatment of rheumatoid arthritis: a preliminary treatment algorithm.* Rheumatology, 2012. **51**(4): p. 600-609.
575. Wijbrandts, C. and P. Tak. *Prediction of response to targeted treatment in rheumatoid arthritis*. in *Mayo Clinic Proceedings*. 2017. Elsevier.
576. Henley, S.S., R.M. Golden, and T.M. Kashner, *Statistical modeling methods: challenges and strategies.* Biostatistics & Epidemiology, 2020. **4**(1): p. 105-139.
577. Van Gestel, A., et al., *Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis: comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism criteria.*

Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 1996. **39**(1): p. 34-40.
578. Felson, D.T., et al., *American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 1995. **38**(6): p. 727-735.
579. Aerts, N.E., et al., *Increased IL-17 production by peripheral T helper cells after tumour necrosis factor blockade in rheumatoid arthritis is accompanied by inhibition of migration-associated chemokine receptor expression.* Rheumatology, 2010. **49**(12): p. 2264-2272.
580. Miao, J., et al., *Circulating Th17 and Th1 cells expressing CD161 are associated with disease activity in rheumatoid arthritis.* Scandinavian Journal of Rheumatology, 2014. **43**(3): p. 194-201.
581. Rakieh, C., et al., *Predicting the development of clinical arthritis in anti-CCP positive individuals with non-specific musculoskeletal symptoms: a prospective observational cohort study.* Annals of the Rheumatic Diseases, 2015. **74**: p. 1659-1666.
582. Daien, C.I., et al., *Regulatory B10 cells are decreased in patients with rheumatoid arthritis and are inversely correlated with disease activity.* Arthritis & rheumatology, 2014. **66**(8): p. 2037-2046.
583. Ma, L., et al., *Reduced numbers of regulatory B cells are negatively correlated with disease activity in patients with new-onset rheumatoid arthritis.* Clinical rheumatology, 2014. **33**(2): p. 187-195.
584. Kane, D., et al., *A prospective, clinical and radiological study of early psoriatic arthritis: an early synovitis clinic experience.* Rheumatology, 2003. **42**(12): p. 1460-1468.
585. McHugh, N., C. Balachrishnan, and S. Jones, *Progression of peripheral joint disease in psoriatic arthritis: a 5-yr prospective study.* Rheumatology, 2003. **42**(6): p. 778-783.
586. Ritchlin, C.T., R.A. Colbert, and D.D. Gladman, *Psoriatic arthritis.* New England Journal of Medicine, 2017. **376**(10): p. 957-970.
587. Ogdie, A. and P. Weiss, *The epidemiology of psoriatic arthritis.* Rheumatic Disease Clinics, 2015. **41**(4): p. 545-568.
588. Mease, P.J. and A.W. Armstrong, *Managing patients with psoriatic disease: the diagnosis and pharmacologic treatment of psoriatic arthritis in patients with psoriasis.* Drugs, 2014. **74**: p. 423-441.
589. Smolen, J., D. Aletaha, and I. McInnes, *Artritis Reumatoide.* The Lancet, 2016. **388**(10055): p. 2023-2038.
590. Singh, J.A., et al., *2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis.* Arthritis & rheumatology, 2016. **68**(1): p. 1-26.
591. Gossec, L., et al., *European League Against Rheumatism recommendations for the management of psoriatic arthritis with pharmacological therapies.* Annals of the rheumatic diseases, 2012. **71**(1): p. 4-12.
592. Kavanaugh, A., et al., *Treatment of psoriatic arthritis in a phase 3 randomised, placebo-controlled trial with apremilast, an oral phosphodiesterase 4 inhibitor.* Annals of the rheumatic diseases, 2014. **73**(6): p. 1020-1026.
593. Mease, P.J., *Measures of psoriatic arthritis: tender and swollen joint assessment, psoriasis area and severity index (PASI), nail psoriasis*

severity index (NAPSI), modified nail psoriasis severity index (MNAPSI), mander/newcastle enthesitis index (MEI), leeds enthesitis index (LEI), spondyloarthritis research consortium of Canada (SPARCC), maastricht ankylosing spondylitis enthesis score (MASES), leeds dactylitis index (LDI), patient global for psoriatic arthritis, dermatology life quality index (DLQI), psoriatic arthritis quality of life (PSAQOL), functional assessment of chronic illness therapy–fatigue (FACIT-F), psoriatic arthritis response criteria (PSARC), psoriatic arthritis joint activity index (PSAJAI), disease activity in psoriatic arthritis (DAPSA), and composite psoriatic disease activity index (CPDAI). Arthritis care & research, 2011. **63**(S11): p. S64-S85.

594. Kavanaugh, A., et al., *Golimumab, a new human tumor necrosis factor α antibody, administered every four weeks as a subcutaneous injection in psoriatic arthritis: twenty-four–week efficacy and safety results of a randomized, placebo-controlled study.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2009. **60**(4): p. 976-986.
595. Mease, P., et al., *Effect of certolizumab pegol on signs and symptoms in patients with psoriatic arthritis: 24-week results of a Phase 3 double-blind randomised placebo-controlled study (RAPID-PsA).* Annals of the rheumatic diseases, 2014. **73**(1): p. 48-55.
596. Mease, P.J., et al., *Etanercept treatment of psoriatic arthritis: safety, efficacy, and effect on disease progression.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2004. **50**(7): p. 2264-2272.
597. Kavanaugh, A., et al., *Golimumab in psoriatic arthritis: one-year clinical efficacy, radiographic, and safety results from a phase III, randomized, placebo-controlled trial.* Arthritis & Rheumatism, 2012. **64**(8): p. 2504-2517.
598. Coates, L.C., et al., *Group for research and assessment of psoriasis and psoriatic arthritis 2015 treatment recommendations for psoriatic arthritis.* Arthritis & rheumatology, 2016. **68**(5): p. 1060-1071.
599. Duffin, K.C., et al., *Association between IL13 polymorphisms and psoriatic arthritis is modified by smoking.* Journal of Investigative Dermatology, 2009. **129**(12): p. 2777-2783.
600. Haroon, M., P. Gallagher, and O. FitzGerald, *Diagnostic delay of more than 6 months contributes to poor radiographic and functional outcome in psoriatic arthritis.* Annals of the rheumatic diseases, 2015. **74**(6): p. 1045-1050.
601. Eder, L., et al., *Factors explaining the discrepancy between physician and patient global assessment of joint and skin disease activity in psoriatic arthritis patients.* Arthritis care & research, 2015. **67**(2): p. 264-272.
602. Maecker, H.T., J.P. McCoy, and R. Nussenblatt, *Standardizing immunophenotyping for the Human Immunology Project.* Nat Rev Immunol, 2012. **12**(3): p. 191-200.
603. Betters, D.M., *Use of flow cytometry in clinical practice.* Journal of the advanced practitioner in oncology, 2015. **6**(5): p. 435.
604. Davis, E.G., M.J. Wilkerson, and B.R. Rush, *Flow cytometry: clinical applications in equine medicine.* Journal of veterinary internal medicine, 2002. **16**(4): p. 404-410.

605. Jahan-Tigh, R.R., et al., *Flow cytometry.* The Journal of investigative dermatology, 2012. **132**(10): p. e1.
606. Kagami, S., et al., *Circulating Th17, Th22, and Th1 cells are increased in psoriasis.* Journal of Investigative Dermatology, 2010. **130**(5): p. 1373-1383.
607. Oldaker, T.A., *Quality control in clinical flow cytometry.* Clinics in laboratory medicine, 2007. **27**(3): p. 671-685.
608. Adan, A., et al., *Flow cytometry: basic principles and applications.* Critical reviews in biotechnology, 2017. **37**(2): p. 163-176.
609. Millán, O. and M. Brunet, *Flow cytometry as platform for biomarker discovery and clinical validation.* General Methods in Biomarker Research and Their Applications; Preedy, VR, Patel, VB, Eds, 2015: p. 141-164.
610. Railean, V. and B. Buszewski, *Flow Cytometry–Sophisticated Tool for Basic Research or/and Routine Diagnosis; Impact of the Complementarity in Both Pre-as Well as Clinical Studies.* Critical Reviews in Analytical Chemistry, 2022: p. 1-23.
611. Bonilla, D.L., G. Reinin, and E. Chua, *Full spectrum flow cytometry as a powerful technology for cancer immunotherapy research.* Frontiers in Molecular Biosciences, 2021. **7**: p. 612801.
612. Alvarez-Barrientos, A., et al., *Applications of flow cytometry to clinical microbiology.* Clinical microbiology reviews, 2000. **13**(2): p. 167-195.
613. Gadalla, R., et al., *Validation of CyTOF against flow cytometry for immunological studies and monitoring of human cancer clinical trials.* Frontiers in oncology, 2019. **9**: p. 415.
614. Hellmich, L., et al., *Flow Cytometry for detection and quantification of micrometastases in sentinel lymph nodes from patients with primary melanoma.* Journal of Surgical Research, 2021. **257**: p. 477-485.
615. Taniguchi, K., et al., *Rapid flow cytometry of gastrointestinal stromal tumours closely matches the modified fletcher classification.* Anticancer research, 2021. **41**(1): p. 131-136.
616. Ponchel, F., A.N. Burska, and S.R. Harrison, *Immunological Biomarkers for Patient Stratification.* 2016: p. 327-361.
617. Dranoff, G., *The therapeutic implications of intratumoral regulatory T cells.* Clinical Cancer Research, 2005. **11**(23): p. 8226-8229.
618. Krijgsman, D., et al., *Characterization of circulating T-, NK-, and NKT cell subsets in patients with colorectal cancer: the peripheral blood immune cell profile.* Cancer Immunology, Immunotherapy, 2019. **68**: p. 1011-1024.
619. Ponchel, F., et al., *Interleukin-7 deficiency in rheumatoid arthritis: consequences for therapy-induced lymphopenia.* Arthritis Res Ther, 2005. **7**(1): p. R80-92.
620. Verburg, R.J., et al., *Outcome of intensive immunosuppression and autologous stem cell transplantation in patients with severe rheumatoid arthritis is associated with the composition of synovial T cell infiltration.* Ann Rheum Dis, 2005. **64**(10): p. 1397-405.
621. Saleem, B., et al., *Patients with RA in remission on TNF blockers: when and in whom can TNF blocker therapy be stopped?* Annals of the Rheumatic Diseases, 2010. **69**(9): p. 1636-42.
622. Davies, J., M. Smyth, and J. Martin, *Automated reticulocyte counting: evaluation of the Coulter® STKS Haematology Analyser reticulocyte counting function.* Clinical & Laboratory Haematology, 1997. **19**(2): p. 89-92.

623. Dunphy, C.H., *Contribution of flow cytometric immunophenotyping to the evaluation of tissues with suspected lymphoma?* Cytometry: The Journal of the International Society for Analytical Cytology, 2000. **42**(5): p. 296-306.
624. Koepke, J.A., *Update on reticulocyte counting.* Laboratory Medicine, 1999. **30**(5): p. 339-343.
625. Stetler-Stevenson, M. and R.C. Braylan. *Flow cytometric analysis of lymphomas and lymphoproliferative disorders*. in *Seminars in hematology*. 2001. Elsevier.
626. Todd, W.M., *Acute myeloid leukemia and related conditions.* Hematology/Oncology Clinics, 2002. **16**(2): p. 301-319.
627. Deptala, A. and S.P. Mayer, *Detection of minimal residual disease.* Methods in Cell Biology, 2001. **64**: p. 385-420.
628. Paietta, E., *Assessing minimal residual disease (MRD) in leukemia: a changing definition and concept?* Bone marrow transplantation, 2002. **29**(6): p. 459-465.
629. de Vries, E., et al., *Flow cytometric immunophenotyping in the diagnosis and follow-up of immunodeficient children.* European journal of pediatrics, 2001. **160**: p. 583-591.
630. Hengel, R.L. and J. Nicholson, *An update on the use of flow cytometry in HIV infection and AIDS.* Clinics in Laboratory Medicine, 2001. **21**(4): p. 841-856.
631. Nicholson, J., *Use of flow cytometry in the evaluation and diagnosis of primary and secondary immunodeficiency diseases.* Archives of pathology & laboratory medicine, 1989. **113**(6): p. 598-605.
632. Hickerson, D.H. and A.P. Bode, *Flow cytometry of platelets for clinical analysis.* Hematology/Oncology Clinics, 2002. **16**(2): p. 421-454.
633. Riley, R.S., *Cellular proliferation markers in the evaluation of human cancer.* Clinics in laboratory medicine, 1992. **12**(2): p. 163-199.
634. Darzynkiewicz, Z., E. Bedner, and P. Smolewski. *Flow cytometry in analysis of cell cycle and apoptosis*. in *Seminars in hematology*. 2001. Elsevier.
635. Darzynkiewicz, Z., P. Smolewski, and E. Bedner, *Use of flow and laser scanning cytometry to study mechanisms regulating cell cycle and controlling cell death.* Clinics in laboratory medicine, 2001. **21**(4): p. 857-873.
636. Nunez, R., *DNA measurement and cell cycle analysis by flow cytometry.* Current issues in molecular biology, 2001. **3**(3): p. 67-70.
637. Bagwell, C.B., et al., *DNA and cell cycle analysis as prognostic indicators in breast tumors revisited.* Clinics in laboratory medicine, 2001. **21**(4): p. 875-95, x.
638. Bagwell, C., et al., *Optimizing flow cytometric DNA ploidy and S-phase fraction as independent prognostic markers for node-negative breast cancer specimens.* Cytometry: The Journal of the International Society for Analytical Cytology, 2001. **46**(3): p. 121-135.
639. Krishan, A., *Monitoring of cellular resistance to cancer chemotherapy: drug retention and efflux.* Methods in Cell Biology, 2001. **64**: p. 193-209.
640. Bogen, K., et al., *Gel microdrop flow cytometry assay for low-dose studies of chemical and radiation cytotoxicity.* Toxicology, 2001. **160**(1-3): p. 5-10.
641. Goguen, B. and N. Kedersha, *Clonogenic cytotoxicity testing by microdrop encapsulation.* Nature, 1993. **363**(6425): p. 189-190.

642. Ellwart, J.W., J.-P. Kremer, and P. Dörmer, *Drug testing in established cell lines by flow cytometric vitality measurements versus clonogenic assay.* Cancer research, 1988. **48**(20): p. 5722-5725.
643. Bleesing, J.J. and T.A. Fleisher. *Cell function-based flow cytometry*. in *Seminars in hematology*. 2001. Elsevier.
644. O'Gorman, M., *Clinically relevant functional flow cytometry assays.* Clinics in laboratory medicine, 2001. **21**(4): p. 779-794.
645. Roesler, J., et al., *Diagnosis of chronic granulomatous disease and of its mode of inheritance by dihydrorhodamine 123 and flow microcytofluorometry.* European journal of pediatrics, 1991. **150**: p. 161-165.
646. O'gorman, M. and V. Corrochano, *Rapid whole-blood flow cytometry assay for diagnosis of chronic granulomatous disease.* Clinical Diagnostic Laboratory Immunology, 1995. **2**(2): p. 227-232.
647. Patton, W., et al., *Assessment of fetal-maternal haemorrhage in mothers with hereditary persistence of fetal haemoglobin.* Journal of clinical pathology, 1990. **43**(9): p. 728-731.
648. Davis, B., et al., *Detection of fetal red cells in fetomaternal hemorrhage using a fetal hemoglobin monoclonal antibody by flow cytometry.* Obstetrical & gynecological survey, 1999. **54**(3): p. 153-154.
649. Nelson, M., et al., *A flow-cytometric equivalent of the Kleihauer test.* Vox Sanguinis, 1998. **75**(3): p. 234-241.
650. Nelson, M., *An overview of the use of flow cytometry in the analysis of mixed red cell populations.* Pathology, 1999. **31**(3): p. 191-198.
651. Fong, E., et al., *Detection of massive transplacental haemorrhage by flow cytometry.* Clinical & Laboratory Haematology, 2000. **22**(6): p. 325-327.
652. Davis, B.H. *Diagnostic advances in defining erythropoietic abnormalities and red blood cell diseases*. in *Seminars in hematology*. 2001. Elsevier.
653. Johnson, P., et al., *Flow cytometry in diagnosis and management of large fetomaternal haemorrhage.* Journal of clinical pathology, 1995. **48**(11): p. 1005-1008.
654. Bromilow, I. and J. Duguid, *Measurement of feto-maternal haemorrhage: a comparative study of three Kleihauer techniques and two flow cytometry methods.* Clinical & Laboratory Haematology, 1997. **19**(2): p. 137-142.
655. Mundee, Y., et al., *Simplified flow cytometric method for fetal hemoglobin containing red blood cells.* Cytometry, 2000. **42**(6): p. 389-393.
656. Horsburgh, T., S. Martin, and A.J. Robson, *The application of flow cytometry to histocompatibility testing.* Transplant Immunology, 2000. **8**(1): p. 3-15.
657. Rebibou, J., et al., *Flow-PRA® evaluation for antibody screening in patients awaiting kidney transplantation.* Transplant immunology, 2000. **8**(2): p. 125-128.
658. Rodrıguez, P.C., et al., *Detection of alloantibodies against non-HLA antigens in kidney transplantation by flow cytometry.* Clinical transplantation, 2000. **14**(5): p. 472-478.
659. Tambur, A.R. and T. Klein, *Flow cytometry application in organ transplantation.* The Israel Medical Association journal: IMAJ, 2000. **2**(4): p. 310-315.
660. Lamb, L.S., *Hematopoietic cellular therapy: implications for the flow cytometry laboratory.* Hematology/Oncology Clinics, 2002. **16**(2): p. 455-476.

661. Storek, J., et al., *Immunity of patients surviving 20 to 30 years after allogeneic or syngeneic bone marrow transplantation.* Blood, The Journal of the American Society of Hematology, 2001. **98**(13): p. 3505-3512.
662. Kishino, K., et al., *Evaluation of engraftment by ABO genotypic analysis of erythroid burst-forming units after bone marrow transplantation.* Leukemia research, 2002. **26**(1): p. 13-17.
663. Winson, M.K. and H.M. Davey, *Flow cytometric analysis of microorganisms.* Methods, 2000. **21**(3): p. 231-240.
664. Shapiro, H.M., *Microbial analysis at the single-cell level: tasks and techniques.* Journal of microbiological methods, 2000. **42**(1): p. 3-16.
665. Steen, H.B., *Flow cytometry of bacteria: glimpses from the past with a view to the future.* Journal of microbiological methods, 2000. **42**(1): p. 65-74.
666. Hřebíček, O., *Applications of flow cytometry in the study of microbial subpopulations.* 2019.
667. Ambriz-Aviña, V., J.A. Contreras-Garduño, and M. Pedraza-Reyes, *Applications of flow cytometry to characterize bacterial physiological responses.* BioMed research international, 2014. **2014**.
668. SA, C., *Binding of ovarian cancer cells to peritoneal mesothlium in vitro is partly mediated by CD44H.* Cancer Res., 1993. **53**: p. 3830-3838.
669. Lessan, K., et al., *CD44 and β1 integrin mediate ovarian carcinoma cell adhesion to peritoneal mesothelial cells.* The American journal of pathology, 1999. **154**(5): p. 1525-1537.
670. Kishikawa, T., et al., *Two distinct patterns of peritoneal involvement shown by in vitro and in vivo ovarian cancer dissemination models.* Invasion & metastasis, 1995. **15**(1-2): p. 11-21.
671. Scotton, C.J., et al., *Epithelial cancer cell migration: a role for chemokine receptors?* Cancer research, 2001. **61**(13): p. 4961-4965.
672. Dong, H.P., et al., *NK-and B-cell infiltration correlates with worse outcome in metastatic ovarian carcinoma.* American journal of clinical pathology, 2006. **125**(3): p. 451-458.
673. Koyama, S., T. Maruyama, and S. Adachi, *Expression of epidermal growth factor receptor and CD44 splicing variants sharing exons 6 and 9 on gastric and esophageal carcinomas: a two-color flow-cytometric analysis.* Journal of cancer research and clinical oncology, 1999. **125**: p. 47-54.
674. Koyama, S., *Enhanced cell surface expression of matrix metalloproteinases and their inhibitors, and tumor-induced host response in progression of human gastric carcinoma.* Digestive diseases and sciences, 2004. **49**: p. 1621-1630.
675. Koyama, S., *Coordinate cell-surface expression of matrix metalloproteinases and their inhibitors on cancer-associated myofibroblasts from malignant ascites in patients with gastric carcinoma.* Journal of cancer research and clinical oncology, 2005. **131**: p. 809-814.
676. Iwata, H., et al., *Flow Cytometric Assay of c-erbB-2 Protein in Fine Needle Aspirates of Fresh and Frozen Human Breast Cancer Tissues.* Breast Cancer, 1996. **3**: p. 111-117.
677. Pachón, G., I. Caragol, and J. Petriz, *Subjectivity and flow cytometric variability.* Nature Reviews Immunology, 2012. **12**(5): p. 396-396.
678. Kalina, T., *Reproducibility of flow cytometry through standardization: opportunities and challenges.* Cytometry Part A, 2020. **97**(2): p. 137-147.

679. Brestoff, J.R. and J.L. Frater, *Contemporary challenges in clinical flow cytometry: small samples, big data, little time.* The journal of applied laboratory medicine, 2022. **7**(4): p. 931-944.
680. Riley, R.S., *Flow cytometry and its applications in hematology and oncology.* Hematology/Oncology Clinics, 2002. **16**(2): p. xi-xii.
681. Brown, M. and C. Wittwer, *Flow cytometry: principles and clinical applications in hematology.* Clinical chemistry, 2000. **46**(8): p. 1221-1229.
682. NHS, S.T.H. *Flow cytometry*. 2023 13 November 2023

]; Available from: https://www.southtees.nhs.uk/services/pathology/haematology-lab/flow-cytometry/.
683. NHS, G.H. *Flow Cytometry Test*. 2023 13 November 2023

]; Available from: https://www.gloshospitals.nhs.uk/our-services/services-we-offer/pathology/haematology/flow-cytometry/
684. NHS, N.u.T.H. *Northern England Haemato-Oncology Diagnostic Service (NEHODS)*. 2023 10 October 2023]; Available from: https://www.newcastlelaboratories.com/lab_service/nehods/.
685. NHS, N.W.L.P. *Immunophenotyping*. 2023 20 October 2023]; Available from: https://www.nwlpathology.nhs.uk/tests-database/immunophenotyping/.
686. Wale, U.H. *Immunophenotyping User Guide*. 2021 20 November 2023]; Available from: https://cavuhb.nhs.wales/files/policies-procedures-and-guidelines/health-and-safety-policies/i-j-health-and-safety/immunophenotyping-users-guide-2016-v2-pdf/.
687. NHS, T.L.T.H. *Flow Cytometry Crossmatch*. 2023 20 November 2023]; Available from: https://www.leedsth.nhs.uk/a-z-of-services/pathology/test-and-tubes/transplant-immunology/flow-cytometry-crossmatch.
688. Correia, R., et al., *A ten-color tube with dried antibody reagents for the screening of hematological malignancies.* International journal of laboratory hematology, 2018. **40**(2): p. 136-143.
689. Hedley, B.D., et al., *Novel lymphocyte screening tube using dried monoclonal antibody reagents.* Cytometry Part B: Clinical Cytometry, 2015. **88**(6): p. 361-370.
690. van der Velden, V.H., et al., *Optimization and testing of dried antibody tube: The EuroFlow LST and PIDOT tubes as examples.* Journal of immunological methods, 2019. **475**: p. 112287.
691. Chan, R.C.F., et al., *Stabilization of pre-optimized multicolor antibody cocktails for flow cytometry applications.* Cytometry Part B: Clinical Cytometry, 2017. **92**(6): p. 508-524.
692. Carulli, G., et al., *Assessment of response to therapy in multiple myeloma by multiparameter flow cytometry. Usefulness of an eight-color single tube with monoclonal antibodies in dried formulation.* La Clinica Terapeutica, 2019. **170**(5): p. e352-e356.
693. Taylor, W., et al., *Classification criteria for psoriatic arthritis: development of new criteria from a large international study.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2006. **54**(8): p. 2665-2673.
694. BD Life Sciences. *BD® Spectrum Viewer | Spectral Analyzer*. 2020 20 November 2021]; Available from: https://www.bdbiosciences.com/en-us/resources/bd-spectrum-viewer.

695. BD Life Sciences – Biosciences. *Fluorochrome/Laser Reference Poster*. 2020 20 November 2021]; Available from: https://www.bdbiosciences.com/content/dam/bdb/marketing-documents/Fluorochrome_Laser_Poster_2.pdf.
696. Streitz, M., et al., *Standardization of whole blood immune phenotype monitoring for clinical trials: panels and methods from the ONE study.* Transplantation research, 2013. **2**: p. 1-15.
697. Bio-Techne. *Flow Cytometry Handbook*. 2024 [cited 2024 22/2/2024]; Available from: https://www.bio-techne.com/resources/literature/flow-cytometry-handbook.
698. Armstrong, R.A., *When to use the B onferroni correction.* Ophthalmic and Physiological Optics, 2014. **34**(5): p. 502-508.
699. Bland, J.M. and D.G. Altman, *Multiple significance tests: the Bonferroni method.* Bmj, 1995. **310**(6973): p. 170.
700. Allison, P.D., *Multiple imputation for missing data: A cautionary tale.* Sociological methods & research, 2000. **28**(3): p. 301-309.
701. Jakobsen, J.C., et al., *When and how should multiple imputation be used for handling missing data in randomised clinical trials–a practical guide with flowcharts.* BMC medical research methodology, 2017. **17**(1): p. 1-10.
702. Tale, A.C., *Multiple imputation for missing data.* SOCIOLOGICAL METHODS & RESEARCH, 2000. **28**(0): p. 301-309.
703. Bartlett, J. and C. Frost, *Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables.* Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology, 2008. **31**(4): p. 466-475.
704. Santegoets, S.J., et al., *Monitoring regulatory T cells in clinical samples: consensus on an essential marker set and gating strategy for regulatory T cell analysis by flow cytometry.* Cancer Immunology, Immunotherapy, 2015. **64**: p. 1271-1286.
705. Koo, T.K. and M.Y. Li, *A guideline of selecting and reporting intraclass correlation coefficients for reliability research.* Journal of chiropractic medicine, 2016. **15**(2): p. 155-163.
706. da Silva Ferreira, V., et al., *The effects of freeze/thawing on the function and phenotype of CD4+ lymphocyte subsets in normal individuals and patients with systemic lupus erythematosus.* Cryobiology, 2015. **71**(3): p. 507-510.
707. Shete, A., et al., *Differential modulation of phenotypic composition of HIV-infected and-uninfected PBMCs during cryopreservation.* Journal of Immunoassay and Immunochemistry, 2013. **34**(4): p. 333-345.
708. Meijerink, M., et al., *Cryopreservation of monocytes or differentiated immature DCs leads to an altered cytokine response to TLR agonists and microbial stimulation.* Journal of immunological methods, 2011. **373**(1-2): p. 136-142.
709. Sattui, S., et al., *Cryopreservation modulates the detection of regulatory T cell markers.* Cytometry Part B: Clinical Cytometry, 2012. **82**(1): p. 54-58.
710. Brooks-Worrell, B., et al., *Comparison of cryopreservation methods on T-cell responses to islet and control antigens from type 1 diabetic patients and controls.* Diabetes/metabolism research and reviews, 2011. **27**(8): p. 737-745.

711. Elkord, E., *Frequency of human T regulatory cells in peripheral blood is significantly reduced by cryopreservation.* Journal of immunological methods, 2009. **347**(1-2): p. 87-90.
712. Costantini, A., et al., *Effects of cryopreservation on lymphocyte immunophenotype and function.* Journal of immunological methods, 2003. **278**(1-2): p. 145-155.
713. Trück, J., et al., *Effect of cryopreservation of peripheral blood mononuclear cells (PBMCs) on the variability of an antigen-specific memory B cell ELISpot.* Human vaccines & immunotherapeutics, 2014. **10**(8): p. 2490-2496.
714. Stevens, V.L., et al., *Cryopreservation of whole blood samples collected in the field for a large epidemiologic study.* Cancer Epidemiology Biomarkers & Prevention, 2007. **16**(10): p. 2160-2163.
715. Hensley-McBain, T., et al., *Optimization of a whole blood phenotyping assay for enumeration of peripheral blood leukocyte populations in multicenter clinical trials.* Journal of immunological methods, 2014. **411**: p. 23-36.
716. Higgins, J., et al., *Effects of lymphocyte isolation and timing of processing on detection of CD127 expression on T cells in human immunodeficiency virus-infected patients.* Clinical and Vaccine Immunology, 2005. **12**(1): p. 228-230.
717. Rabin, T., et al., *Association of circulating lymphocyte subsets with response to IL17i and TNFi in axial spondyloarthritis.* Clinical and experimental rheumatology, 2023. **41**(3): p. 694-703.
718. Johnson, S., V. Nguyen, and D. Coder, *Assessment of cell viability.* Current protocols in cytometry, 2013. **64**(1): p. 9.2. 1-9.2. 26.
719. Shenkin, M., R. Babu, and R. Maiese, *Accurate assessment of cell count and viability with a flow cytometer.* Cytometry Part B: Clinical Cytometry, 2007. **72**(5): p. 427-432.
720. Kummrow, A., et al., *Quantitative assessment of cell viability based on flow cytometry and microscopy.* Cytometry Part A, 2013. **83**(2): p. 197-204.
721. Medzhitov, R., *Origin and physiological roles of inflammation.* Nature, 2008. **454**(7203): p. 428-435.
722. Rui, L., *Energy metabolism in the liver.* Comprehensive physiology, 2014. **4**(1): p. 177.
723. Lauer, F.T., J.L. Denson, and S.W. Burchiel, *Isolation, cryopreservation, and immunophenotyping of human peripheral blood mononuclear cells.* Current protocols in toxicology, 2017. **74**(1): p. 18.20. 1-18.20. 16.
724. Lemieux, J., et al., *A global look into human T cell subsets before and after cryopreservation using multiparametric flow cytometry and two-dimensional visualization analysis.* Journal of immunological methods, 2016. **434**: p. 73-82.
725. Anioke, I., et al., *Lymphocyte subset phenotyping for the prediction of progression to inflammatory arthritis in anti-citrullinated-peptide antibody-positive at-risk individuals.* Rheumatology, 2023: p. kead466.
726. Alivernini, S., et al., *The B side of rheumatoid arthritis pathogenesis.* Pharmacological Research, 2019. **149**: p. 104465.
727. Fathollahi, A., et al., *The role of NK cells in rheumatoid arthritis.* Inflammation Research, 2021. **70**(10): p. 1063-1073.

728. Gizinski, A.M. and D.A. Fox, *T cell subsets and their role in the pathogenesis of rheumatic disease.* Current Opinion in Rheumatology, 2014. **26**(2).
729. Tu, J., et al., *A tale of two immune cells in rheumatoid arthritis: the crosstalk between macrophages and T cells in the synovium.* Frontiers in Immunology, 2021. **12**: p. 2359.
730. Mankia, K. and P. Emery, *Preclinical Rheumatoid Arthritis: Progress Toward Prevention.* Arthritis & Rheumatology (Hoboken, NJ), 2016. **68**(4): p. 779-788.
731. Gelman, A., et al., *Multiple imputation for model checking: Completed-data plots with missing and latent data.* Biometrics, 2005. **61**(1): p. 74-85.
732. Abayomi, K., A. Gelman, and M. Levy, *Diagnostics for multivariate imputations.* Journal of the Royal Statistical Society Series C: Applied Statistics, 2008. **57**(3): p. 273-291.
733. He, Y. and A.M. Zaslavsky, *Diagnosing imputation models by applying target analyses to posterior replicates of completed data.* Statistics in medicine, 2012. **31**(1): p. 1-18.
734. Nguyen, C.D., J.B. Carlin, and K.J. Lee, *Diagnosing problems with imputation models using the Kolmogorov-Smirnov test: a simulation study.* BMC medical research methodology, 2013. **13**: p. 1-9.
735. Gul, H.L., et al., *Defining remission in rheumatoid arthritis: does it matter to the patient? A comparison of multi-dimensional remission criteria and patient reported outcomes.* Rheumatology, 2020. **59**(3): p. 613-621.
736. Solomon, D.H., et al., *Treat to target in rheumatoid arthritis: fact, fiction or hypothesis?* Arthritis & rheumatology (Hoboken, NJ), 2014. **66**(4): p. 775.
737. Smolen, J.S., et al., *Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force.* Annals of the rheumatic diseases, 2016. **75**(1): p. 3-15.
738. Genolini, C., et al., *kml and kml3d: R packages to cluster longitudinal data.* Journal of statistical software, 2015. **65**: p. 1-34.
739. Kryszczuk, K. and P. Hurley. *Estimation of the number of clusters using multiple clustering validity indices*. in *Multiple Classifier Systems: 9th International Workshop, MCS 2010, Cairo, Egypt, April 7-9, 2010. Proceedings 9*. 2010. Springer.
740. Tighe, R.M., et al., *Improving the quality and reproducibility of flow cytometry in the lung. An official American Thoracic Society workshop report.* American journal of respiratory cell and molecular biology, 2019. **61**(2): p. 150-161.
741. Diks, A.M., et al., *Impact of blood storage and sample handling on quality of high dimensional flow cytometric data in multicenter clinical research.* Journal of immunological methods, 2019. **475**: p. 112616.
742. Aimee Tran, D.L., Mamatha Bharadwaj1, Stephanie Widmann, Colin Williams, Suraj Saksena, John Archdeacon and Elena Afonina. *Development of a Dried-down, Multicolor Reagent*

Solution for Enhanced Flow-cytometric Applications. 2019 17 November 2022].
743. Li, Y., E. Mateu, and I. Díaz, *Impact of cryopreservation on viability, phenotype, and functionality of porcine PBMC.* Frontiers in immunology, 2021. **12**: p. 765667.

744. Elliott, G.D., S. Wang, and B.J. Fuller, *Cryoprotectants: A review of the actions and applications of cryoprotective solutes that modulate cell recovery from ultra-low temperatures.* Cryobiology, 2017. **76**: p. 74-91.
745. Heng, B.C., et al., *Loss of viability during freeze–thaw of intact and adherent human embryonic stem cells with conventional slow-cooling protocols is predominantly due to apoptosis rather than cellular necrosis.* Journal of biomedical science, 2006. **13**: p. 433-445.
746. Gómez-Mora, E., et al., *Impact of long-term cryopreservation on blood immune cell markers in myalgic encephalomyelitis/chronic fatigue syndrome: implications for biomarker discovery.* Frontiers in immunology, 2020: p. 2974.
747. Draxler, D., et al., *A flowcytometric analysis to efficiently quantify multiple innate immune cells and T Cell subsets in human blood.* Cytometry Part A, 2017. **91**(4): p. 336-350.
748. Maecker, H.T., J.P. McCoy, and R. Nussenblatt, *Standardizing immunophenotyping for the human immunology project.* Nature Reviews Immunology, 2012. **12**(3): p. 191-200.
749. Rasmussen, S.M., et al., *Stable phenotype of B-cell subsets following cryopreservation and thawing of normal human lymphocytes stored in a tissue biobank.* Cytometry Part B: Clinical Cytometry, 2015. **88**(1): p. 40-49.
750. Rundgren, I.M., et al., *Standardization of sampling and sample preparation for analysis of human monocyte subsets in peripheral blood.* Journal of immunological methods, 2018. **461**: p. 53-62.
751. Thyagarajan, B., et al., *Effect of delayed cell processing and cryopreservation on immunophenotyping in multicenter population studies.* Journal of immunological methods, 2018. **463**: p. 61-70.
752. Ticha, O., L. Moos, and I. Bekeredjian-Ding, *Effects of long-term cryopreservation of PBMC on recovery of B cell subpopulations.* Journal of immunological methods, 2021. **495**: p. 113081.
753. Hønge, B.L., et al., *Optimizing recovery of frozen human peripheral blood mononuclear cells for flow cytometry.* PloS one, 2017. **12**(11): p. e0187440.
754. Björkström, N.K., et al., *Characterization of natural killer cell phenotype and function during recurrent human HSV-2 infection.* PLoS One, 2011. **6**(11): p. e27664.
755. Reimann, K.A., et al., *Preservation of lymphocyte immunophenotype and proliferative responses in cryopreserved peripheral blood mononuclear cells from human immunodeficiency virus type 1-infected donors: implications for multicenter clinical trials.* Clinical Diagnostic Laboratory Immunology, 2000. **7**(3): p. 352-359.
756. Kreher, C.R., et al., *CD4+ and CD8+ cells in cryopreserved human PBMC maintain full functionality in cytokine ELISPOT assays.* Journal of immunological methods, 2003. **278**(1-2): p. 79-93.
757. Bull, M., et al., *Defining blood processing parameters for optimal detection of cryopreserved antigen-specific responses for HIV vaccine trials.* Journal of immunological methods, 2007. **322**(1-2): p. 57-69.
758. Mallone, R., et al., *Isolation and preservation of peripheral blood mononuclear cells for analysis of islet antigen-reactive T cell responses: position statement of the T-Cell Workshop Committee of the Immunology*

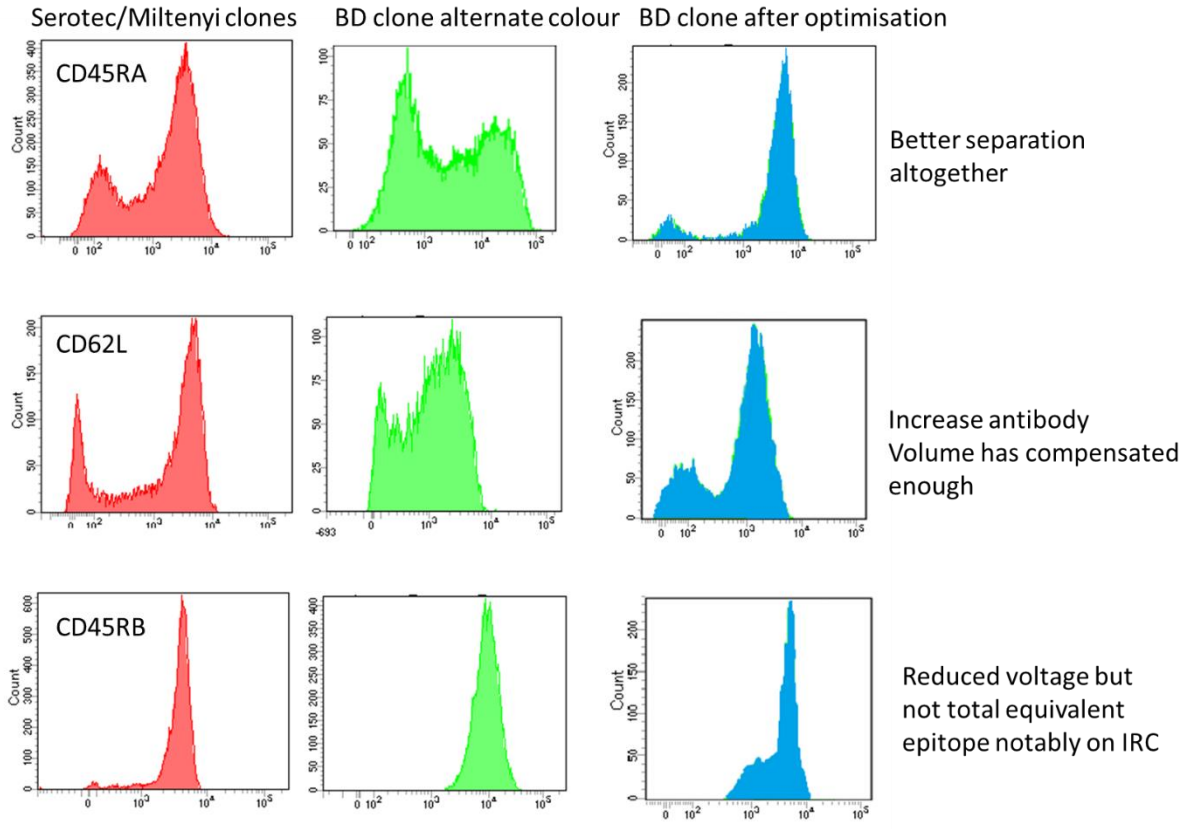
of Diabetes Society. Clinical & Experimental Immunology, 2011. **163**(1): p. 33-49.
759. Tompa, A., A. Nilsson-Bowers, and M. Faresjö, *Subsets of CD4+, CD8+, and CD25hi lymphocytes are in general not influenced by isolation and long-term cryopreservation.* The Journal of Immunology, 2018. **201**(6): p. 1799-1809.
760. Bos, W.H., et al., *Arthritis development in patients with arthralgia is strongly associated with anti-citrullinated protein antibody status: a prospective cohort study.* Annals of the Rheumatic Diseases, 2010. **69**: p. 490-494.
761. Katchamart, W., et al., *Associations of rheumatoid factor and anti-citrullinated peptide antibody with disease progression and treatment outcomes in patients with rheumatoid arthritis.* Rheumatology international, 2015. **35**: p. 1693-1699.
762. Avouac, J., L. Gossec, and M. Dougados, *Diagnostic and predictive value of anti-CCP (cyclic citrullinated protein) antibodies in rheumatoid arthritis: a systematic literature review.* Annals of the rheumatic diseases, 2006.
763. Bizzaro, N., et al., *Anti-cyclic citrullinated peptide antibody titer predicts time to rheumatoid arthritis onset in patients with undifferentiated arthritis: results from a 2-year prospective study.* Arthritis research & therapy, 2013. **15**: p. 1-9.
764. van Der Helm-van Mil, A.H., et al., *Validation of a prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: moving toward individualized treatment decision-making.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2008. **58**(8): p. 2241-2247.
765. van der Linden, M.P., et al., *Value of anti–modified citrullinated vimentin and third-generation anti–cyclic citrullinated peptide compared with second-generation anti–cyclic citrullinated peptide and rheumatoid factor in predicting disease outcome in undifferentiated arthritis and rheumatoid arthritis.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 2009. **60**(8): p. 2232-2241.
766. Bos, W.H., et al., *Arthritis development in patients with arthralgia is strongly associated with anti-citrullinated protein antibody status: a prospective cohort study.* Annals of the rheumatic diseases, 2010. **69**(3): p. 490-494.
767. Rosser, E.C., et al., *Regulatory B cells are induced by gut microbiota–driven interleukin-1β and interleukin-6 production.* Nature medicine, 2014. **20**(11): p. 1334-1339.
768. Yoshizaki, A., et al., *Regulatory B cells control T-cell autoimmunity through IL-21-dependent cognate interactions.* Nature, 2012. **491**(7423): p. 264-268.
769. Carter, N.A., et al., *Mice lacking endogenous IL-10–producing regulatory B cells develop exacerbated disease and present with an increased frequency of Th1/Th17 but a decrease in regulatory T cells.* The Journal of Immunology, 2011. **186**(10): p. 5569-5579.
770. Bankó, Z., et al., *Induction and differentiation of IL-10–producing regulatory B cells from healthy blood donors and rheumatoid arthritis patients.* The Journal of Immunology, 2017. **198**(4): p. 1512-1520.

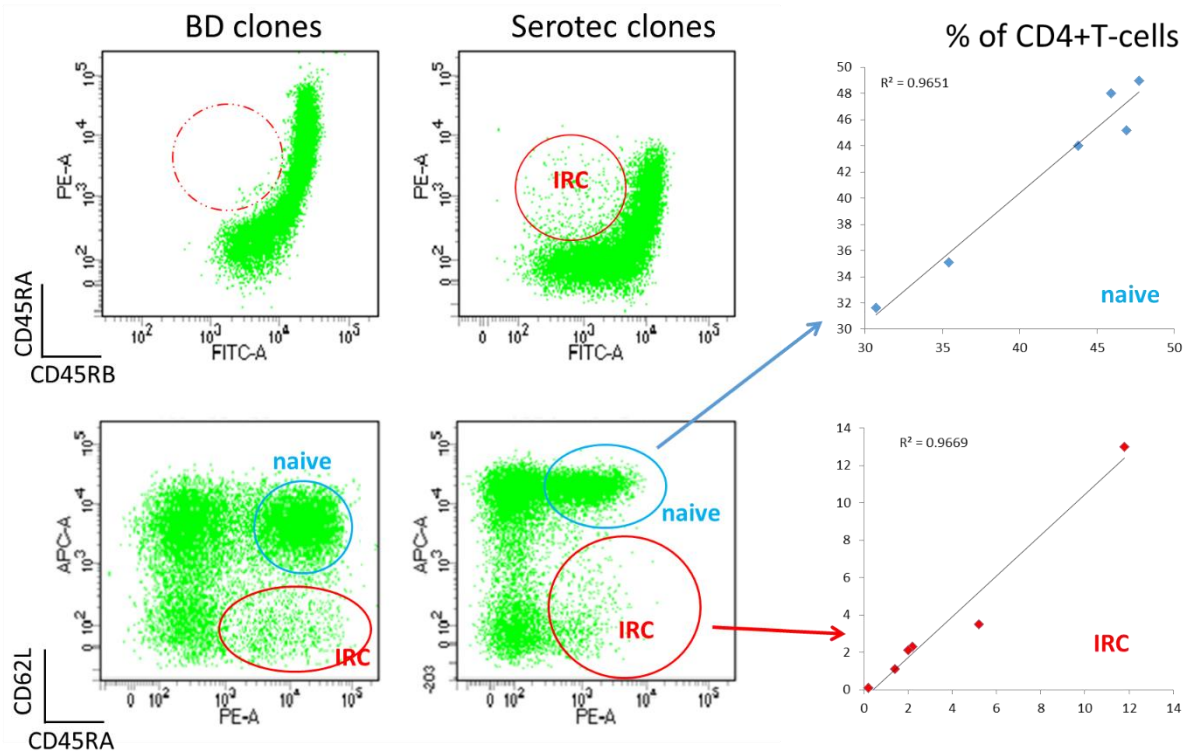
771. Mauri, C., *CD19 CD24 hi CD38 hi B Cells Exhibit Regulatory Capacity in Healthy Individuals but Are Functionally Impaired in Systemic Lupus Erythematosus Patients.* Immunity, 2010. **32**: p. 129-140.
772. Churchman, S.M., et al., *Modulation of peripheral T-cell function by interleukin-7 in rheumatoid arthritis.* Arthritis research & therapy, 2014. **16**(6): p. 1-13.
773. Gul, H., K. Harnden, and B. Saleem. *Defining the Optimal Strategies for Achieving Drug-Free Remission in Rheumatoid Arthritis: A Narrative Review*. in *Healthcare*. 2021. Multidisciplinary Digital Publishing Institute.
774. Tian, Y., et al., *Unique phenotypes and clonal expansions of human CD4 effector memory T cells re-expressing CD45RA.* Nature communications, 2017. **8**(1): p. 1-13.
775. Larbi, A. and T. Fulop, *From "truly naïve" to "exhausted senescent" T cells: when markers predict functionality.* Cytometry Part A, 2014. **85**(1): p. 25-35.
776. Lugli, E., et al., *Superior T memory stem cell persistence supports long-lived T cell memory.* The Journal of clinical investigation, 2013. **123**(2).
777. Maldonado, A., et al., *Decreased effector memory CD45RA+ CD62L- CD8+ T cells and increased central memory CD45RA-CD62L+ CD8+ T cells in peripheral blood of rheumatoid arthritis patients.* Arthritis Res Ther, 2003. **5**: p. 1-6.
778. Arroyo-Villa, I., et al., *Frequency of Th17 CD4+ T cells in early rheumatoid arthritis: a marker of anti-CCP seropositivity.* 2012.
779. Su, R., et al., *Altered distribution of circulating T follicular helper-like cell subsets in rheumatoid arthritis patients.* Frontiers in Medicine, 2021. **8**: p. 690100.
780. Symmons, D., et al., *Lymphopenia in rheumatoid arthritis.* Journal of the Royal Society of Medicine, 1989. **82**(8): p. 462-463.
781. Ponchel, F., et al., *Interleukin-7 deficiency in rheumatoid arthritis: consequences for therapy-induced lymphopenia.* Arthritis Res Ther, 2004. **7**: p. 1-13.
782. Pawłowska, J., et al., *Different distribution of CD4 and CD8 T cells in synovial membrane and peripheral blood of rheumatoid arthritis and osteoarthritis patients.* Folia Histochemica et Cytobiologica, 2009. **47**(4): p. 627-632.
783. Leheita, O., et al., *Lymphocytes subsets in osteoarthritis versus rheumatoid arthritis.* The Egyptian Journal of Immunology, 2005. **12**(2): p. 113-124.
784. Zhao, J., et al., *Apoptosis, Autophagy, NETosis, Necroptosis, and Pyroptosis Mediated Programmed Cell Death as Targets for Innovative Therapy in Rheumatoid Arthritis.* Frontiers in immunology, 2021: p. 5652.
785. Aramaki, T., et al., *A significantly impaired natural killer cell activity due to a low activity on a per-cell basis in rheumatoid arthritis.* Modern Rheumatology, 2009. **19**(3): p. 245-252.
786. Al-Laith, M., et al., *Arthritis prevention in the pre-clinical phase of RA with abatacept (the APIPPRA study): a multi-centre, randomised, double-blind, parallel-group, placebo-controlled clinical trial protocol.* Trials, 2019. **20**(1): p. 1-15.
787. Goëb, V., et al., *Progression to rheumatoid arthritis in early inflammatory arthritis is associated with low IL-7 serum levels.* Annals of the Rheumatic Diseases, 2013. **72**(6): p. 1032-1036.

788. EuroFlow Consortium, E.-F.L. *Standardized flow cytometry protocols*. 2024 [cited 2023 12/10/2023]; Available from: https://euroflow.org/protocols/.
789. Coates, L.C. and P.S. Helliwell, *Methotrexate efficacy in the tight control in psoriatic arthritis study.* The Journal of rheumatology, 2016. **43**(2): p. 356-361.
790. Helliwell, P.S., W.J. Taylor, and C.S. Group, *Treatment of psoriatic arthritis and rheumatoid arthritis with disease modifying drugs--comparison of drugs and adverse reactions.* The Journal of Rheumatology, 2008. **35**(3): p. 472-476.
791. Felquer, M.L.A., et al., *Drug therapies for peripheral joint disease in psoriatic arthritis: a systematic review.* The Journal of rheumatology, 2014. **41**(11): p. 2277-2285.
792. Conigliaro, P., et al., *Restoration of peripheral blood natural killer and B cell levels in patients affected by rheumatoid and psoriatic arthritis during etanercept treatment.* Clinical & Experimental Immunology, 2014. **177**(1): p. 234-243.
793. Lobão, B., et al., *From PsO to PsA: the role of TRM and Tregs in psoriatic disease, a systematic review of the literature.* Frontiers in Medicine, 2024. **11**: p. 1346757.
794. Wang, J., et al., *The numbers of peripheral regulatory T cells are reduced in patients with psoriatic arthritis and are restored by low-dose interleukin-2.* Therapeutic advances in chronic disease, 2020. **11**: p. 2040622320916014.
795. Duquenne, L., et al., *Predicting inflammatory arthritis in at-risk persons: development of scores for risk stratification.* Annals of Internal Medicine, 2023. **176**(8): p. 1027-1036.
796. England, B.R., et al., *Increased cardiovascular risk in rheumatoid arthritis: mechanisms and implications.* Bmj, 2018. **361**.
797. Fragoulis, G.E., I. Panayotidis, and E. Nikiphorou, *Cardiovascular risk in rheumatoid arthritis and mechanistic links: from pathophysiology to treatment.* Current Vascular Pharmacology, 2020. **18**(5): p. 431-446.
798. Lu, B., et al., *Being overweight or obese and risk of developing rheumatoid arthritis among women: a prospective cohort study.* Annals of the rheumatic diseases, 2014. **73**(11): p. 1914-1922.
799. Hohmeier, H., et al., *Inflammatory mechanisms in diabetes: lessons from the β-cell.* International Journal of Obesity, 2003. **27**(3): p. S12-S16.
800. Shi, J., et al., *Anti-carbamylated protein (anti-CarP) antibodies precede the onset of rheumatoid arthritis.* Annals of the rheumatic diseases, 2014. **73**(4): p. 780-783.
801. Kudo-Tanaka, E., et al., *Autoantibodies to cyclic citrullinated peptide 2 (CCP2) are superior to other potential diagnostic biomarkers for predicting rheumatoid arthritis in early undifferentiated arthritis.* Clinical rheumatology, 2007. **26**: p. 1627-1633.
802. Greenland, S. and W.D. Finkle, *A critical look at methods for handling missing covariates in epidemiologic regression analyses.* American journal of epidemiology, 1995. **142**(12): p. 1255-1264.
803. Hayati Rezvan, P., K.J. Lee, and J.A. Simpson, *The rise of multiple imputation: a review of the reporting and implementation of the method in medical research.* BMC medical research methodology, 2015. **15**: p. 1-14.

804. Schober, P. and T.R. Vetter, *Missing data and imputation methods.* Anesthesia & Analgesia, 2020. **131**(5): p. 1419-1420.
805. Sterne, J.A., et al., *Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls.* Bmj, 2009. **338**.
806. Zhang, Y., et al., *A systematic survey of the methods literature on the reporting quality and optimal methods of handling participants with missing outcome data for continuous outcomes in randomized controlled trials.* Journal of clinical epidemiology, 2017. **88**: p. 67-80.
807. Jørgensen, A.W., et al., *Comparison of results from different imputation techniques for missing data from an anti-obesity drug trial.* PLoS One, 2014. **9**(11): p. e111964.
808. Shah, A.D., et al., *Low eosinophil and low lymphocyte counts and the incidence of 12 cardiovascular diseases: a CALIBER cohort study.* Open heart, 2016. **3**(2): p. e000477.
809. Fohner, A.E., et al., *Association of peripheral lymphocyte subsets with cognitive decline and dementia: the cardiovascular health study.* Journal of Alzheimer's Disease, 2022. **88**(1): p. 7.
810. Liu, Z.-Q., M.-Y. Lu, and B. Liu, *Circulating CD56+ NKG2D+ NK cells and postoperative fertility in ovarian endometrioma.* Scientific Reports, 2020. **10**(1): p. 18598.
811. Donders, A.R.T., et al., *A gentle introduction to imputation of missing values.* Journal of clinical epidemiology, 2006. **59**(10): p. 1087-1091.
812. Schafer, J.L. and M.K. Olsen, *Multiple imputation for multivariate missing-data problems: A data analyst's perspective.* Multivariate behavioral research, 1998. **33**(4): p. 545-571.
813. Graham, J.W., *Missing data: Analysis and design*. 2012: Springer Science & Business Media.
814. Collins, L.M., J.L. Schafer, and C.-M. Kam, *A comparison of inclusive and restrictive strategies in modern missing data procedures.* Psychological methods, 2001. **6**(4): p. 330.
815. Graham, J.W., et al., *Analysis with missing data in prevention research.* 1997.
816. Rubin, D.B., *Multiple imputation after 18+ years.* Journal of the American statistical Association, 1996. **91**(434): p. 473-489.

Appendix A


Serotec/Miltenyi clones
BD clone alternate colour
BD clone after optimisation
CD45RA
CD62L
CD45RB
Count
Better separation altogether
Increase antibody Volume has compensated enough
Reduced voltage but not total equivalent epitope notably on IRC


BD clones
Serotec clones
% of CD4+T-cells
PE-A
FITC-A
CD45RA
CD45RB
IRC
naive
R² = 0.9651
APC-A
PE-A
CD62L
CD45RA
R² = 0.9669

Appendix B Ethical Approval

The Leeds Teaching Hospitals NHS Trust

Date: 26th February 2020
Our Ref: RR09/9134

LTHT Resourcing Service
Trust Headquarters
St James's University Hospital
Beckett Street
Leeds
LS9 7TF

Direct Line (0113) 2065980
www.leedsth.nhs.uk

Private and Confidential
Mr Anioke Innocent
C/o LIRMM
Wellcome Trust Brenner Building
St.James`s University Hospital
Beckett Street
Leeds Teaching Hospitals
LS9 7TF

Dear Mr Innocent,

Letter of Access for Research - Project Title: RADAR - Rheumatoid Arthritis Disease Research

This letter confirms your right of access to conduct research through The Leeds Teaching Hospitals NHS Trust for the purpose and on the terms and conditions set out below. This right of access commences on **25th February 2020** and ends on **24th February 2023** unless terminated earlier in accordance with the clauses below.

You have a right of access to conduct such research as confirmed in writing in the letter of permission for research from this NHS organisation. Please note that you cannot start the research until the Principal Investigator for the research project has received a letter from us giving permission to conduct the project.

The information supplied about your role in research at The Leeds Teaching Hospitals NHS Trust has been reviewed and you do not require an honorary research contract with this NHS organisation. We are satisfied that such pre-engagement checks as we consider necessary have been carried out.

You are considered to be a legal visitor to The Leeds Teaching Hospitals NHS Trust premises. You are not entitled to any form of payment or access to other benefits provided by this NHS organisation to employees and this letter does not give rise to any other relationship between you and this NHS organisation, in particular that of an employee.

While undertaking research through The Leeds Teaching Hospitals NHS Trust, you will remain accountable to your employer **University of Leeds** but you are required to follow the reasonable instructions of **Sinisa Savic** in this NHS organisation or those given on her/his behalf in relation to the terms of this right of access.

You must act in accordance with The Leeds Teaching Hospitals NHS Trust policies and procedures, which are available to you upon request, and the Research Governance Framework.

Where any third party claim is made, whether or not legal proceedings are issued, arising out of or in connection with your right of access, you are required to co-operate fully with any

investigation by this NHS organisation in connection with any such claim and to give all such assistance as may reasonably be required regarding the conduct of any legal proceedings.

You are required to co-operate with The Leeds Teaching Hospitals NHS Trust in discharging its duties under the Health and Safety at Work etc Act 1974 and other health and safety legislation and to take reasonable care for the health and safety of yourself and others while on The Leeds Teaching Hospitals NHS Trust premises. You must observe the same standards of care and propriety in dealing with patients, staff, visitors, equipment and premises as is expected of any other contract holder and you must act appropriately, responsibly and professionally at all times.

You are required to ensure that all information regarding patients or staff remains secure and *strictly confidential* at all times. You must ensure that you understand and comply with the requirements of the NHS Confidentiality Code of Practice (http://www.dh.gov.uk/assetRoot/04/06/92/54/04069254.pdf) and the Data Protection Act 2018. Furthermore you should be aware that under the Act, unauthorised disclosure of information is an offence and such disclosures may lead to prosecution.

You should ensure that, where you are issued with an identity or security card, a bleep number, email or library account, keys or protective clothing, these are returned upon termination of this arrangement. Please also ensure that while on the premises you wear your ID badge at all times, or are able to prove your identity if challenged. Please note that this NHS organisation accepts no responsibility for damage to or loss of personal property.

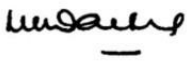
We may terminate your right to attend at any time either by giving seven days' written notice to you or immediately without any notice if you are in breach of any of the terms or conditions described in this letter or if you commit any act that we reasonably consider to amount to serious misconduct or to be disruptive and/or prejudicial to the interests and/or business of this NHS organisation or if you are convicted of any criminal offence. Where required by law, your HEI employer will initiate your Independent Safeguarding Authority (ISA) registration, and thereafter, will continue to monitor your ISA registration status via the on-line ISA service. Should you cease to be ISA-registered, this letter of access is immediately terminated. Your employer will immediately withdraw you from undertaking this or any other regulated activity. You MUST stop undertaking any regulated activity.

Your substantive employer is responsible for your conduct during this research project and may in the circumstances described above instigate disciplinary action against you.

The Leeds Teaching Hospitals NHS Trust will not indemnify you against any liability incurred as a result of any breach of confidentiality or breach of the Data Protection Act 2018. Any breach of the Data Protection Act 2018 may result in legal action against you and/or your substantive employer.

If your current role or involvement in research changes, or any of the information provided in your Research Passport changes, you must inform your employer through their normal procedures. You must also inform your nominated manager in this NHS organisation.

Yours sincerely



Mandy Walker
Head of Resourcing

Appendix C Ethical Approval

NHS

Leeds (West) Research Ethics Committee
A/B Floor, Old Site
Leeds General Infirmary
Great George Street
Leeds
LS1 3EX

Telephone: 0113 3923181
Facsimile: 0113 392 2863

07 November 2006

Prof Paul Emery
Professor of Rheumatology/Consultant Rheumatologist
university of Leeds
Rhematology, 2nd Floor
Chapel Allerton Hospital
Leeds
LS7 4SA

Dear Prof Emery

Full title of study: **Coordinated Programme to Prevent Arthritis:Can We Identify Arthritis at a Pre-Clinical Stage?**

REC reference number: **06/Q1205/169**

Thank you for your letter of 26 October 2006, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised.

Ethical review of research sites

The Committee has designated this study as exempt from site-specific assessment (SSA. There is no requirement for [other] Local Research Ethics Committees to be informed or for site-specific assessment to be carried out at each site.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Application		19 July 2006
Investigator CV		
Protocol	3.5	20 October 2006

Covering Letter		19 July 2006
Letter from Sponsor		26 July 2006
Peer Review		12 February 2006
Questionnaire: Employment Questions	3.3 (in protocol)	
Questionnaire: Health Assessment	3.3 (in protocol)	
Questionnaire: for Anti-CCP negative patients	3.3 (in protocol)	18 July 2006
GP/Consultant Information Sheets	1.0	17 August 2006
Participant Information Sheet: Parts 1 & 2	2.3	20 October 2006
Participant Information Sheet: Part 3	1.1	20 October 2006
Participant Consent Form: Parts 1 & 2	2.3	20 October 2006
Participant Consent Form: Part 3	1.1	20 October 2006
Response to Request for Further Information		26 October 2006
Response to Request for Further Information		13 September 2006

Research governance approval

You should arrange for the R&D department at all relevant NHS care organisations to be notified that the research will be taking place, and provide a copy of the REC application, the protocol and this letter.

All researchers and research collaborators who will be participating in the research must obtain final research governance approval before commencing any research procedures. Where a substantive contract is not held with the care organisation, it may be necessary for an honorary contract to be issued before approval for the research can be given.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

06/Q1205/169	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project

Yours sincerely

Laura Sawiuk
REC Co-ordinator
On Behalf of
DR JOHN PUNTIS
Chair

Email: laura.sawiuk@leedsth.nhs.uk

Enclosures:

Standard approval conditions

Appendix D Ethical Approval

ICON

NHS
National Research Ethics Service

Leeds (West) Research Ethics Committee
First Floor
Millside
Mill Pond Lane
Leeds
LS6 4EP

Tel: 0113 305 0116

24 August 2010

Professor Paul Emery
LEEDS TEACHING HOSPITALS NHS TRUST
Department of Rheumatology. Second Floor.
Chapel Allerton Hospital, Chapeltown Road
Leeds
LS7 4SA

Dear Professor Emery

Study title:	**Inflammatory arthritis disease continuum longitudinal study**
REC reference:	**09/H1307/98**
Protocol number:	**RR09/9134**
Amendment number:	**1**
Amendment date:	**05 August 2010**

The above amendment was reviewed at the meeting of the Sub-Committee held on 17 August 2010.

Ethical opinion

Favourable Opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the provision that the following points be adhered to:

1) The word 'research' should be added to the invitation letter, after the phrase "potentially suitable for the ICON Disease continuum".

2) The invitation letter should come from (and be signed by) the research team – not the Consultant in charge of the patient's care.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Sample request form		
Protocol	3.0	02 August 2010
Letter of invitation to participant	1.0	02 August 2010
Participant Information Sheet: Information Sheet and Consent Form	3.0	02 August 2010

Participant Information Sheet: Information Sheet and Consent Form	3.0	02 August 2010
Protocol	3.0	02 August 2010
Notice of Substantial Amendment (non-CTIMPs)		05 August 2010

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

09/H1307/98:	**Please quote this number on all correspondence**

Yours sincerely

Claire Kelly
Committee Assistant Co-ordinator

E-mail: Claire.kelly@leedspft.nhs.uk

Copy to: *Mrs Rachel de Souza*

Appendix E SOP for PMBC Isolation

LIRMM Chapel Allerton Standard Operating Procedure	Title	SOP for sample processing: PBMC isolation using Leucosep™ tubes				
	Version	1.0	Date	3.2.20	SOP_ID	SOP 30S

STANDARD OPERATING PROCEDURE FOR THE

ISOLATION and FREEZING of PBMC USING LEUCOSEP™ TUBES

Details:

1.1.1.1

SOP Number: **SOP 30S**

Effective Date: **Immediately following approval**

Approval:

Version No: of the SOP being approved.	Name of person approving this SOP.	Date	Signature of the person approving this SOP.
1.0	Dr Elena Jones	08/02/2020	

Distribution & Storage:

Location of Document

Paper: Approved SOP folder - CAH LAB Office

Electronic: N:\Faculty-of-Medicine-and-Health\IMMECR\Musculo- Skeletal\Lab Samples\SOPs\NEW ‘

1. **Purpose and scope**

This SOP provides instructions for the isolation of PBMC (Peripheral Blood Mononuclear Cells) using pre-filled Leucosep^{TM} tubes at LIRMM Chapel Allerton.

This SOP must be followed to ensure a high standard of separation and storage of biological samples as the way a sample is collected, transported, fractionated and stored will affect sample integrity which is extremely important to the final laboratory or research results.

2. **Health and safety**

2.1 **All employees should make themselves aware of any health and safety issues related to the use of Chemical and biological hazards and demonstrate adequate training has been received. Employees are responsible for ensuring the health and safety of themselves and others in the workplace.**

2.2 Biological samples may represent an infection risk; therefore appropriate protective equipment should be worn. Lab coat and gloves should be worn whenever handling biological samples. Dispose of all biological solid waste in biological hazardous waste container.

3. Detailed description of the method.

3.1 Blood processed in the CAH Lab for PBMC can be collected in Lithium Heparin (green top) or K2ETDA (purple top) tubes.

The blood collection must be carried out by a suitably trained registered nurse, doctor or phlebotomist and should follow site-specific practices.

Blood tubes should be labelled either with the sample barcode or where handwritten, labelling should be legible and in permanent ink.

In the case of any very specific collections the procedure will be agreed with the lab to maintain required conditions through the study.

CHAPTER 8 MATERIALS AND EQUIPMENT

3.2.1 Equipment:

Laminar Flow Cabinet

ALC PK 130R centrifuge

Freezer -150C and -80C

Pipetteboy

Materials:

Sterile 50ml centrifuge tubes and sterile 15ml conical tubes

Sterile 2ml tubes (e.g. Greiner BIO-ONE 123263)

Inserts for tubes e.g. Sarstedt 65.386

Sterile 10ml, 5ml and 1ml pipettes

Sterile pastettes

Storage box suitable for -150^{C} freezer

Mr Frosty or Cool Cell

Pre-filled 50ml Leucosep^{TM} tubes eg Greiner Bio-one 227 288

3.2.2 Reagents

Sterile Phosphate Buffered Saline (PBS) (e.g. Gibco 18912-014)

Fetal calf serum (FCS) heat-inactivated eg Gibco 16140063

DMSO eg Sigma D2650

Isopropanol (for Mr Frosty)

3.3 PROCEDURE for PBMC separation from blood samples using Leucosep™ tubes (carry out procedure in a laminar flow cabinet)

3.3.1 If any of the density gradient medium has risen above the porous barrier in the Leucosep™ tube during transportation then a 1minute spin at 1000g is required before use.

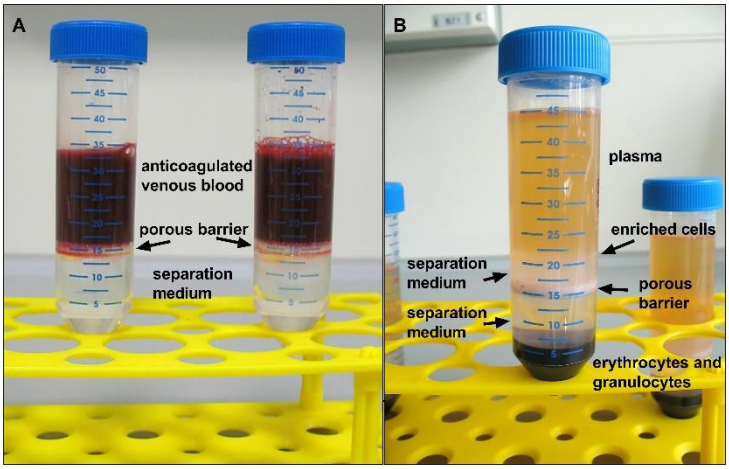
3.3.2 Dilute the blood sample in an equal volume of sterile PBS

E.g. if you have 15 ml of blood add 15 ml of sterile PBS

3.3.3 Pour the blood/PBS mixture gently on top of the barrier in the tube.

3.3.4 Centrifuge at 2500 rpm/900g for 10 minutes with no chiller and no brake (programme 4 on ALC PK 130R).

3.3.5 After centrifuging the PBMC layer appears as a cloudy layer between the plasma and separation medium.



3.3.6 Harvest the PBMC layer using a sterile pastette or by pouring the liquid phase above the barrier directly into a 50ml centrifuge tube. Use a small volume of PBS to wash the Leucosep™ tube and add this to the cells. Make up the volume to 40ml with PBS.

3.3.7 Centrifuge at 1200rpm /250g for 10 minutes (programme 5 on ALC PK 130R).

3.3.8 Pour off the supernatant, resuspend the pellet and make up the volume to 40 ml with sterile PBS.

Wash by centrifuging at 1200 rpm/250g for 10 minutes.

3.3.8 Repeat washing step using a 15ml tube and 10ml of PBS.

Proceed to either 3.4 or 3.5

3.4 TO SAVE THE CELLS AS A DRY PELLET (eg FOR DNA ANALYSIS)

3.4.1 Pour off supernatant from 3.3.8, resuspend pellet and transfer to a 2ml tube

3.4.2 Make up the volume to 2ml using PBS

3.4.3 Spin at 7500rpm/5000g for 10 minutes in Hettich Micro 20 centrifuge.

3.4.4 Pour off supernatant and blot on tissue paper until all the liquid is removed.

3.4.5 Freeze at -80C.

3.5 TO SAVE THE CELLS IN A VIABLE STATE

3.5.1 Pour off the supernatant from 3.3.8. The lymphocytes are frozen in foetal calf serum (FCS) containing 10% dimethylsulfoxide (DMSO) in a volume of 1 ml per vial. The FCS/DMSO must be kept chilled in order to reduce cell damage.

3.5.2 Add cells to 1ml freezing mixture (90% FBS + 10% DMSO) in 2ml vial and place in a Mr Frosty or Cool Cell in -80C freezer overnight before placing in -150C freezer for storage.

NB if required cells may be counted using a haemocytometer or Bio Rad TC20 prior to freezing (see SOP 23E).

4 Roles and responsibilities

This SOP applies to all CAH Lab staff collecting research samples. Where applicable this SOP must be used in conjunction with HTA Codes of Practice and all other relevant University and, where appropriate, local University Health and Safety policies and SOPS.

4.1 The clinical staff are responsible for:

- completion of the appropriate Research sample request form for the particular study with correct and complete information in order to correct collection and recording of the sample.

- data should include donor, study and time of collection.

- signing the appropriate sample collection form by which they confirm that the patient has signed appropriate consent for the sample to be collected and used as part of the study/research purpose.

The phlebotomists are responsible for:

- identifying the patient against the details on the form before specimen collection, filling in collection time on the form. It is the responsibility of the person taking the sample to identify the patient, label the sample and ensure that the information supplied on the request form and sample are accurate and match in each case.

4.2 The Laboratory personnel are responsible for:

- Working in accordance with this SOP.

- Checking Research Blood Request Form and sample quality.

- Recording relevant information about the sample condition into the study sample log.

- Tracing missing information if reasonably possible.

- Processing samples that have been collected accompanied by a complete and signed sample collection form.

- Responding to queries.

5. References

1. Human Tissue Act 2004

2. Human Tissue Authority, Codes of practice 9: Research

Appendix F SOP FOR B cell panel

Routine standard flow-cytometry technique followed by the NHS-lab

(A) B-cell panel SOP

procedure or methodology
Before handling biological samples and chemical contained within this SOP familiarize yourself with the COSHH and **Risk Assessments** and take the necessary precautions required.
The sample is evaluated for risk status in the sample reception area of Room 07 09 154. All request cards/ forms must be stamped and booking in form should be completed at the sample reception desk. Once completed, it should be attached to the request card and put on the clipboard located next to sample reception tray. The original request card and the booking in form should have a telepath bar code sticker attached to them and the rest of the stickers should be attached to the sample bag. A sticker should be attached to the sample tube when the sample is taken from the bag.
If the staining procedure is not performed straight away the sample should remain in its bag and stored in the high risk sample reception 20°C incubator in Room 07 09 161.

1. Labelled test tubes with the patient initials and the name of the panel and placed them in a rack.
2. Labelled 2 test tubes for each patient initials and the name of the panel and placed them in a rack.

3. Add 300 µl of patient's sample (whole blood) into one of the test tube and add 3 ml of PBS.

4. Vortex and then centrifuge tubes at 1500rpm for 5 minutes; make sure tubes are balanced.

5. Aspirate supernatant and resuspend the cells in 3 mL of 1X PBS.

6. Repeat step 4 for the second wash.

7. After the final centrifuge, aspirate supernatant; add 1500 µl of 1% PBS into the cell pellet and vortex.

8. Add antibodies as listed below into the stain tube, ensure that the antibodies are pipette to the bottom of the tube and none is left on the side of the tube.

1 µl CD19-BV421 (Biolegend cat # 302234)	4 µl CD10-APC (BD cat # 332777)
2 µl IgD-V500 (BD cat # 561490)	10 µl CD27-PE (BD cat # 555441)
2 µl CD38-PE-Cy7 (BD cat # 335825)	10 µl CD24-FITC (BD cat # 555427)
2 µl IgM-PerCP-Cy5.5 (BD cat # 561285)	

Change pipette for each antibody.

9. Vortex the patient cell pellet (washed blood) and add 300 µL into the appropriate stain tube. Ensure blood sample and the antibodies are properly mixed. Ensure that the blood is pipette to the bottom of the tube and none is left on the side of the tube as it will not be incubated with the monoclonal antibodies and in analysis will appear as a false negative population. These volumes are critical and should be measured accurately as it will affect the results.

10. When dealing with more than one sample check the name on the **EDTA** tube matches those on the FACS Flow tubes.

11. Mixed gently and incubate for 20 minutes at room temperature protecting samples from the light during this time.

12. Add 5mL of lysing solution (3/4 full) into each tube, vortex and incubate for 20 minutes at room temperature.

13. Centrifuge tubes at 1500 rpm for 5 minutes; make sure tubes are balanced (symmetrical).

14. Decant supernatant and resuspend cells.

15. Add 3 mL of 1% PBS into each test tube and centrifuge at 1500rpm for 5 minutes

16. Decant supernatant and repeat washing procedure (as above).

17. Decant supernatant and resuspend in PBS + 0.5% formaldehyde approximately 5 drops.

18. Mixed solution and run on Flow Cytometer. Analyse within 24 hours on the FACSCanto Flow Cytometer using the B-cell template and acquiring 30,000 B cell events/samples.

19. Sample can be analysed immediately on the flow cytometer or analysed the following day. For next day analysis stored stained tubes in the fridge at 4ºC, protectd from light.

N.B the Flow cytometer must be set up following SOP's- Operation of FACSCanto Flow Cytometer and Calibration of Flow Cytometer (QC)

Appendix G SOP for LS panel

(B)

procedure or methodology

Before handling biological samples and chemical contained within this SOP familiarize yourself with the COSHH and **Risk Assessments** and take the necessary precautions required.

The sample is evaluated for risk status in the sample reception area of Room 07 09 154. All request cards/ forms must be stamped and booking in form should be completed at the sample reception desk. Once completed, it should be attached to the request card and put on the clipboard located next to sample reception tray. The original request card and the booking in form should have a telepath bar code sticker attached to them and the rest of the stickers should be attached to the sample bag. A sticker should be attached to the sample tube when the sample is taken from the bag.

If the staining procedure is not performed straight away the sample should remain in its bag and stored in the high risk sample reception 20°C incubator in Room 07 09 161.

1. Labelled test tubes with the patient initials and the name of the panel and placed them in a rack.

2. Add antibodies as listed below into the stain tube, ensure that the antibodies are pipette to the bottom of the tube and none is left on the side of the tube. Change pipette tip for each antibody.

1 µl CD19-BV421 (BD Cat # 562440)	10 µl CD3-FITC (BD Cat # 555332)
2 µl CD4-V500 (BD Cat # 560768)	15 µl CD56-PE (BD Cat # 555516)
3 µl CD16-PE-Cy7(BD Cat # 557744)	15 µl CD14-APC (BD Cat # 555399)
5 µl CD8-PerCP5 (BD Cat # 560662)	

3. Invert (mix) patient blood and add 300 µL, ensure blood sample and the antibodies are properly mixed. Ensure that the blood is pipette to the bottom of the tube and none is left on the side of the tube as it will not be incubated with the monoclonal antibodies and in analysis will appear as a false negative population. These volumes are critical and should be measured accurately as it will affect the results.

4. When dealing with more than one sample check the name on the **EDTA** tube matches those on the FACS Flow tubes.

5. Mixed gently and incubate for 20 minutes at room temperature (in the dark) protecting samples from the light during this time.

6. Add pre-lysing solution (3/4 full) to each tube, vortex and incubate for 15 minutes at room temperature.

7. Centrifuge tubes at 1500 rpm for 5 minutes; make sure tubes are balanced (symmetrical).

8. Decant supernatant and resuspend cells.

9. Add 3 mL of 1% PBS into each test tube and centrifuge at 1500rpm for 5 minutes.

10. Decant supernatant and repeat washing procedure (as above).

11. Decant supernatant and resuspend in PBS + 0.5% formaldehyde approximately 5 drops.

12. Mixed solution and run on Flow Cytometer. Analyse within 24 hours on the FACSCanto Flow Cytometer using the **Cell Count template** and acquiring **100,000 lymphocyte events/samples**.

13. Sample can be analysed immediately on the flow cytometer or analysed the following day. For next day analysis stored stained tubes in the fridge at 4ºC.

N.B the Flow cytomer must be set up following SOP's- Operation of FACSCanto Flow Cytometer and Calibration of Flow Cytometer (QC)
See SOP…….. For retrieving Cell Count template, recording results for exporting and importing dat

Appendix H T-cell subsets panel SOP

The Leeds Teaching Hospitals NHS NHS Trust **Pathology Directorate**	**F Examination Process** **F2 Examination procedures** Printed on : 12/06/2024 Page 425 of 3
Specialist Laboratory Medicine	
Section : Transplant and Cellular Immunology	

Title: Staining Whole Blood Samples for Regulatory T-Cell (T-Reg) Analysis on the FACSCanto Flow Cytometer

Site/Area of application	Transplant Immunology, Cellular Immunology
Index code	
Superseded documents	
Implementation date of this version	July 2013
Approver of content of SOP	Dr Clive Carter/ Zainab Kabba
Reason for change	
Keywords for search on EQMS	

Signature

1. CLINICAL RELEVANCE/PURPOSE OF PROCEDURE

Regulatory T (T-Reg) cells are an important subset of lymphocytes and are believed to be essential for maintaining peripheral tolerance, preventing autoimmune diseases and limiting chronic inflammatory diseases. However, they may also have limit beneficial responses by suppressing immunity and limiting anti-tumour responses. The phenotypes of human T-reg cells is still incompletely understood, with those that express CD3, CD4, CD25 and FoxP3 being the most understood. Identification and enumeration of these cells can be helpful in certain pathological states.

2. PRINCIPLE OF PROCEDURE

When monoclonal antibodies are added to whole blood, the fluorochrome labelled antibodies bind specifically on the surface of the leucocytes, which enable enumeration of the sub- populations of white cells using FACSCanto flow cytometer. As the transcription factor Foxp3 is localised within the nuclear compartment of cells, this protocol depends on cell surface combined with intracellular staining.

3. PERSONNEL / TRAINING REQUIREMENTS

Only those individuals who have received specific training from Zainab Kabba or Dr Clive Carter are permitted to perform the procedures outlined in this SOP. Trainee Biomedical Scientists and Trainee Clinical Scientists and above are permitted to perform these tasks without direct supervision.

4. SPECIMEN REQUIREMENTS

Minimum 2 mL **EDTA** blood, to reach the laboratory within 24 hours of collection, and stored at room temperature. Other anticoagulant blood samples are **NOT** acceptable.

5. EQUIPMENT

- Pipette tips polypropylene 5-100 µL yellow, pack of 1000 (Regional Supplies Dept).
- Pipette tips polypropylene 200-1000 µL blue, pack of 100 (Regional Supplies Dept).
- Pastettes (Regional Supplies Dept).
- Eppendorf pipette 0.5- 10 µL
- Eppendorf pipette 200- 1000 µL
- Gloves, disposable (Regional Supplies Dept).
- Goggles, coverall (BDH)
- Laboratory coat
- Laboratory Booking in Forms
- FACSCanto Flow Cytometer (Becton Dickinson)
- FACS Flow Tubes
- FACS falcon tubes

6. HEALTH AND SAFETY/RISK ASSESSMENT

See COSHH and procedure risk assessments

7. REAGENTS

- FACSFlow solution, FASC Shut-down solution, FACS Cleaning solution (Becton Dickinson).
- FACS Flow Tubes (Becton Dickinson)

- Monoclonal antibodies (Becton Dickinson) see below for details.

- Lysing solution (Becton Dickinson), dilute to 1/10 with distilled water.

- 1% PBS / FBS (wash buffer) - made with 500 mL PBS and 5 mL of foetal bovine serum (FBS Invitrogen).

- **Permeabilization/Wash buffer-** (by mixing **1 part** of permeabilization concentrate and **9 parts** of the ddH_2O. (eg. 1 mL solution A and 9 mL ddH2O).

- Permeabilization buffer (**buffer C**) - by mixing **1 part** of fix/Permeabilization concentrate and **3 parts** of the fix/permeabilization diluent. (980 µL of buffer A + 20 µL of perm (buffer B).

- 0.5% Formaldehyde (fix solution): made with 1.35 mL 37% formaldehyde (BDH) and 100 mL PBS.

8. CALIBRATION

See **SOP……..**

9. QUALITY CONTROL

Several quality control issues are relevant to this assay:-
A check list is attached to the booking in form and must be fully signed before a result is reported on Telepath. Also see ‘Acceptance Criteria’ Section.

10. COMPUTER / TELEPATH CODES

N/A

11. PROCEDURE OR METHODOLOGY

Before handling biological samples and chemical contained within this SOP familiarize
yourself with the **COSHH** and **Risk Assessments** and take the necessary precautions
required.
The sample is evaluated for risk status in the sample reception area of Room 07 09 154.
All request cards/ forms must be stamped and booking in form should be completed at
the sample reception desk. Once completed, it should be attached to the request card
and put on the clipboard located next to sample reception tray. The original request card
and the booking in form should have a telepath bar code sticker attached to them and

the rest of the stickers should be attached to the sample bag. A sticker should be
attached to the sample tube when the sample is taken from the bag.

If the staining procedure is not performed straight away the sample should remain in its
bag and stored in the high risk sample reception 20°C incubator in Room 07 09 161.

1. Labelled test tubes with the patient initials and the name of the panel and placed them in a rack. Add antibodies as listed below into the stain tube, ensure that the antibodies are pipette to the bottom of the tube and none is left on the side of the tube. Change pipette tip for each antibody.

2 µl HLA-DR-APC-H7 (BD cat # 561358)	5 µl CD127- PerCP-Cy5.5 (BD cat # 560551)
2 µl CD25-PE-Cy7 (BD cat # 335824)	5 µl CD45RA-PE(Serotec cat # MCA88PE)
2 µl CD4-BV421(BD cat# 562424)	5 µl CD62L-APC (Miltenyi cat #130-091-755
4 µl CD3-V500 (BD cat # 561416)	

2. Mix the blood sample by inversion and add 300 µL/tube, ensuring that the blood and the antibodies are properly mixed. Also check that the blood is pipetted to the bottom of the tube and none is left on the side of the tube as it will not be incubated with the monoclonal antibodies and in analysis will appear as a false negative population. These volumes are critical and should be measured accurately as it will affect the results.

3. Mixed gently and incubate for 20 minutes at room temperature (in the dark) protecting samples from the light during this time.

4. Add lysing solution (4.5ml) to each tube, vortex and incubate for 10 minutes at room temperature.

5. Centrifuge tubes at 1500 rpm for 5 minutes; make sure tubes are balanced (symmetrical).

6. Decant supernatant and resuspend cells by vortexing.
7. Add 4 mL of 1% PBS into each test tube and centrifuge at 1500rpm for 5 minutes

8. Decant supernatant and repeat washing procedure (as above).

9. Decant supernatant and resuspend cells by vortexing.

10. Add 2 mL of diluted buffer A (Fix buffer) into each test tube and incubate for 10 minutes.

11. Centrifuge tubes at 1500rmp for 5 minutes and decant supernatant. Resuspend cell pellet by vortexing

12. Add 0.5 mL of sample C (Permeabilization buffer), mix by votexing and incubate for 30 minutes at RT in the dark.

13. Add 3mL of PBS into each test tube and wash (see steps 7-9).

14. Add 2 µL of FOX- P3-Alexa488 antibody (BD cat # 561181), mix and incubate for 30 minutes in the dark.

15. Wash cells (repeat steps 7-9).

16. Decant supernatant and resuspend in PBS + 0.5% formaldehyde (approximately 200uL).

    Analyse within 24 hours on the FACSCanto Flow Cytometer using the **T-Reg template** (**T-Reg V2 15/05/15**) and acquiring **30,000 events CD4+ T cells/samples.** The Flow cytotomer must be set up following SOP's- Operation of FACSCanto Flow Cytometer and Calibration of Flow Cytometer (QC)

17. Sample can be analysed immediately on the flow cytometer or analysed the following day. For next day analysis stored stained tubes in the fridge at 4ºC, protected from light.

18. Results reported on Telepath are % of $CD4^+$ T cells positive for CD25 and Foxp3.

See SOP Title: Rheumatology Project analysis using Diva 7 Software to retrieve **T-Reg template**, recording results and exporting and importing data

12. UNCERTAINTY OF MEASUREMENT

See acceptance criteria

13. REFERENCE RANGE / ACTION LIMITS

Age related reference ranges for each subset are printed on the telepath report. Advice may be sought from the consultant Immunologist/rheumatologist if values fall unexpectedly outside the reference ranges.

14. REFERENCES

See monoclonal antibody package inserts (Becton Dickinson)

15.

Appendix I Summary of T and Treg panel SOP

T-cell panel and Treg Panel Protocol

mAb/Sample	Tube 1(TCP)	Tube 2(Treg)
1. mAb: CD4-BV421 CD62L-APC CD3-V500 CD45RA-PE CD45RB-FITC CD25-PE-CY7 CD127-Per-Cy5.5	 2µl 2µl 4µl 5µl 7.5µl - -	 2µl 2µl 4µl - - 2µl 5µl
2. Blood(200-300µl)	250µl	250µl
✓ Incubate 20mins @ rtm ✓ Half fill with FACS lysing buffer ✓ Vortex ✓ Top up with FACS lysing buffer ✓ Incubate 10mins @rtm ✓ Spin and wash with PBS@1500rpm 5mins, repeat until clear		
3. FACS buffer 4. formaldehyde	300µl 4 drops Vortex->Flow	-
5. Solution A	-	2ml
✓ Vorex ✓ Incubate 10mins @ rtm ✓ Spin/wash		
6. Solution C	-	0.5ml
		✓ Vortex ✓ Incubate 30mins @ rtm ✓ Top with PBS and wash/spin x2
7. Foxp3-Alexa488	-	3µl Incubate in dark for 30 minuets Wash twice
8. FACS buffer 9. formaldehyde	-	300µl 5 drops Vortex->Flow

Preparation of FoxP3 Buffer (for permeabilisation)
Label two tubes A and C
Pipette 0.5ml FOXP3 buffer [redacted] into A to make 1 in 10 dilution using H2O (NB:per 2 patients)
Take 980µl of A into C (NB: Per 2 patients)
Add 20µl FOXP3 (small brown container) into C (NB: per 2 patients)

FoxP3: H20
0.5ml : 4.5ml → 0.5/5 = 5/50 =1/10 (i.e. 1:9)

Appendix J protocol for preparing PBMC from frozen cells

Materials

PBS containing DNAase (aliquots done. 100ul/50ml….2ul/ml) keep constant throughout PBS and FACS

(to make DNase form lypholysed …2.75ml PBS/ vial)

Blocking buffer…100uL human IgG, 300uL mouse serum, 1100uL FACs buffer in store at 4^{0}C

PBS solution= 1 table of PBS + 200ml distilled H2O

FACS Buffer (500ml) =
1. 500ml PBS
2. 0.5g Bovine Saline Albumin (BSA)
3. 5ml of 1% Sodium Azide (SA)
4. 200µL of 0.5M EDTA solution
5. filter using the filter unit (must be done in the hood)

Protocol.

1. Thaw cells quickly in a water bath at 37C, till frozen solid can be removed from vial by tipping.
2. Decant into 9mls PBS containing DNase at RT. Mix by inverting x3
3. Spin at 500g, 5mins 12^{0}C
4. Decant supernatant, resuspend pellet in 1ml PBS, top up to 10ml PBS. Mix by inverting.
5. Spin at 500g , 5mins 12^{0}C
6. Decant supernatant, dry pellet by inverting on towel for few seconds.
7. Remove and residual liquid with a pipette, if necessary do not disturb the pellet.
8. Resuspend pellet in blocking buffer(BB) (volume dependent on size of pellet)* usually 50-100uL
9. Aliquot 50ul into each tube.
10. Add Abs according to vol required (see above antibody panel table)
11. Leave at 4^{0}C for 30mins
12. Wash with about 2mls or more of PBS containing DNase, spin as above
13. Decant, resuspend in PBS, spin as above
14. Decant, resuspend cell in about 300uL FACS buffer
15. Acquire in flow cytometry

Keep away from direct light

Buffer

Appendix K

Buffers

Lysis buffer: Ammonium chloride buffer 1X (Ammonium Chloride (A4514-Sigma) 8.99g + Potassium Bicarbonate (P/5080/53- Fisher Scientific) 1g + Na EDTA0.5 M 200ul all top up to one litre with ddH_2O) **Keep for 5 days maximum**

Blocking buffer: (Mouse serum (M5905-Sigma) 300ul + Human IgG (I2511-Sigma) 100ul +FACS buffer 1100 ul).

FACS buffer: 500ml (0.1% BSA in PBS (0.5g) + 0.01% Sodium Azide (1ml of 5% stock solution) + 200ul Na EDTA 0.5 M)

Fix solution: 500ml of PBS + 20 gms of paraformaldehyde powder (Sigma Cat# P6148) heat on hotplate magnetic stirrer the fume hood at 60°C till fluid is clear. Cool to RT. Then filter to store at 4°C (Fridge) for 6

Fix/Permeabilization buffer: (by mixing **1 part** of fix/Permeabilization concentrate (ebioscience) cat# 00-5123-43 and **3 parts** of the fix/permeabilization diluent (ebioscience) cat# 00-5123-56).
Prepare **200ul** per patient sample

Permeabilization/Wash buffer: (by mixing **1 part** of permeabilization concentrate (ebioscience) cat# 00-8333-56 (ebioscience) and **9 parts** of the ddH_2O). Prepare 1ml/patient sample

Book centrifuge for 2 hours

Collect bloods

ICON samples post IDEA/EMPIRE keep 2ml for Ann Morgan Group

Lyse RBCs- 1st lysis

Add 3mls blood to 42ml lysis buffer - mix by inversion.

Leave for 10 minutes at RT

Centrifuge

LIDS ON CENTRIFUGE POTS- make sure pots are balanced (symmetrical)

500rcf (*), acc8 + decel8, 8-10mins, RT

Clean up any blood spill in centrifuge buckets using 2% Trigene spray / alcohol

Decant the supernatant

Reddish colour. Decant into Distel.

Should leave a white pellet

Re-suspend pellet

1ml **PBS (phosphate buffer saline)** –

re-suspend using P1000 pipette

0.5ml (500µl) PBS for each 50ml falcon tube: transfer the re-suspended pellet from one falcon tube to the other (using pipette) so that all of the cells from that one sample you separated in the first instance (step 2) are in the same tube

Leave for 5 mins

Leaving for longer will not hurt- just enhances lysis

Second lysis- Add 40ml Lysis buffer

Mix by inversion

Incubate for **5-10** minutes

Nb adding more lysis buffer will NOT hurt (<50ml- fill falcon tube)

Centrifuge for 2nd lysis
500rcf(*), acc8, decel8, 8-10mins, RT
Use LIDS for centrifuge pots
Clean any spills of blood with 2% Trigene

Decant the supernatant
pale pinkish lysis supernatant. Decant into Trigene container.

Aspirate reddish supernatant remaining above pellet
Using pipette P1000

Re-suspend pellet
0.5ml (500µl) **PBS** – using pipette P1000

Add *10ml* PBS, re-suspend and filter
Filter into 50ml falcon tube, with 10ml pipette + pipette aid + 70µm strainer

RINSE strainer with *10ml* PBS
Final vol 20ml PBS

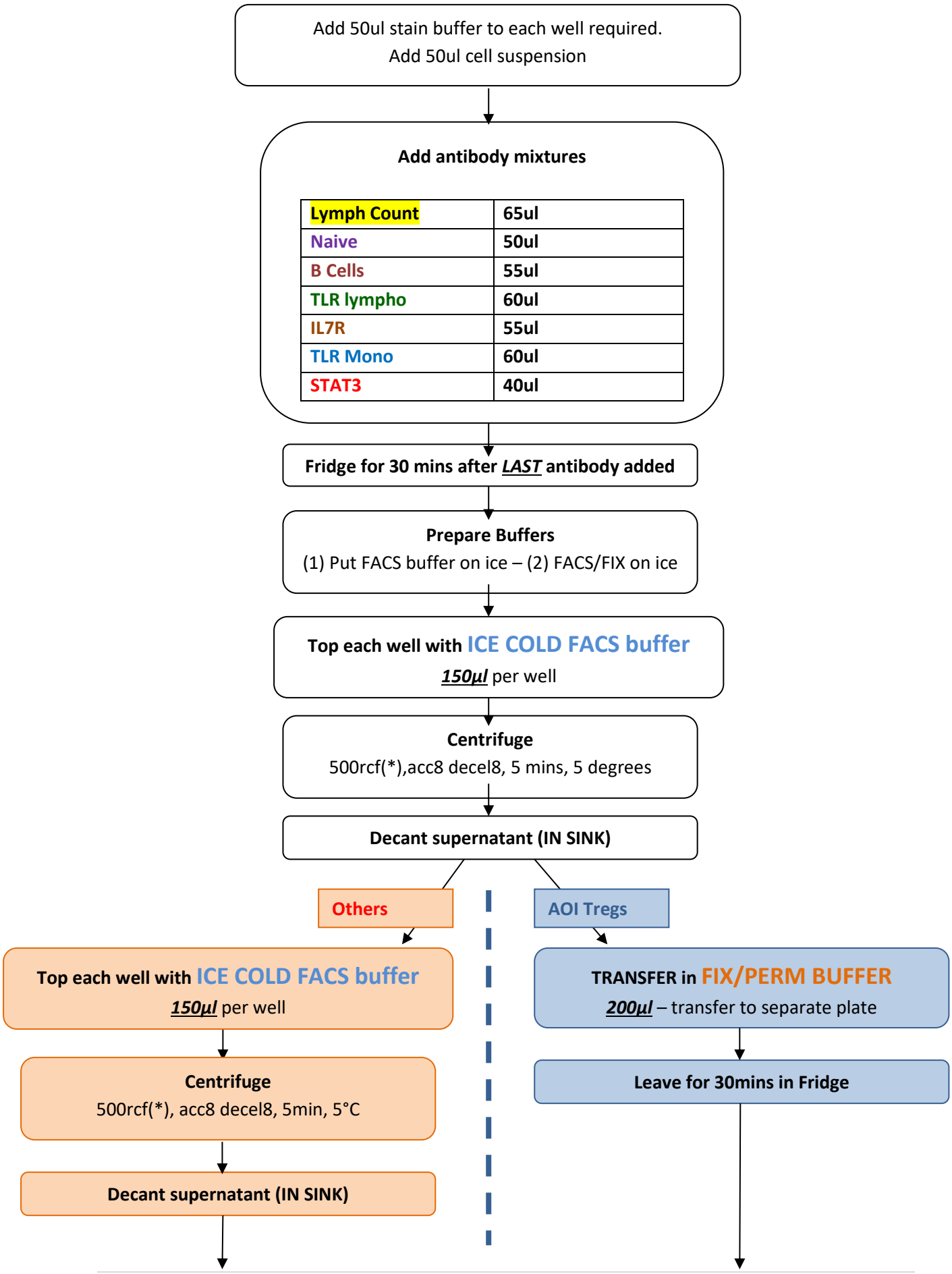
If required you should prepare fix/perm and perm/wash buffers during the centrifuge steps

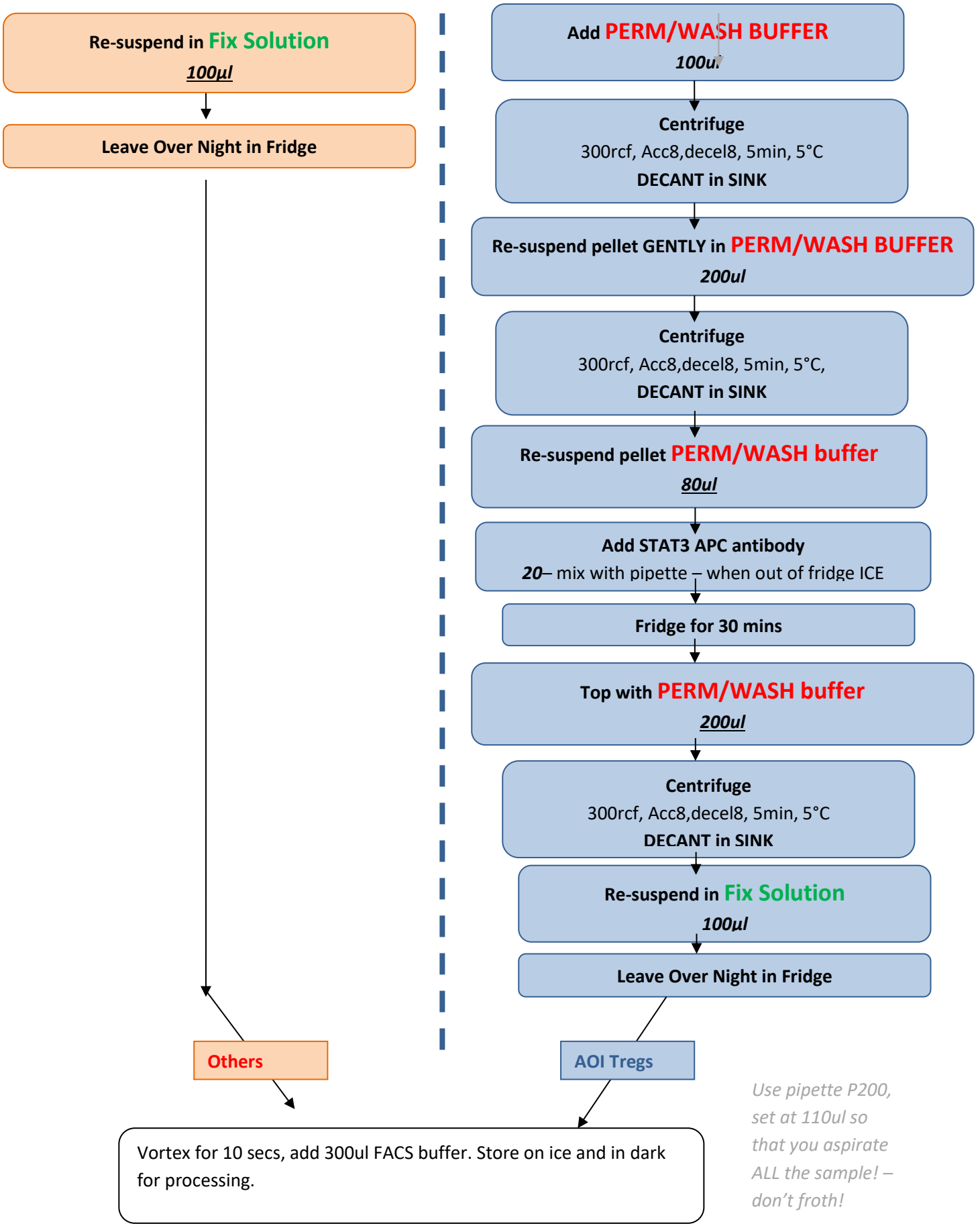
Centrifuge
500rcf(*),8acc, 8decel, 8-10mins, 5°C,
Use LIDS for centrifuge pots
Clean any spills of blood with 2% Trigene spray

Decant the supernatant
Decant into Trigene container

Aspirate reddish supernatant remaining above pellet
Using pipette P1000

RE-suspend in BLOCKING BUFFER
Buffer required = ***no. of wells x 50µl***


Add 50ul stain buffer to each well required.
Add 50ul cell suspension
Add antibody mixtures
Lymph Count 65ul
Naive 50ul
B Cells 55ul
TLR lympho 60ul
IL7R 55ul
TLR Mono 60ul
STAT3 40ul
Fridge for 30 mins after LAST antibody added
Prepare Buffers
(1) Put FACS buffer on ice – (2) FACS/FIX on ice
Top each well with ICE COLD FACS buffer
150µl per well
Centrifuge
500rcf(*),acc8 decel8, 5 mins, 5 degrees
Decant supernatant (IN SINK)
Others
AOI Tregs
Top each well with ICE COLD FACS buffer
150µl per well
TRANSFER in FIX/PERM BUFFER
200µl – transfer to separate plate
Centrifuge
500rcf(*), acc8 decel8, 5min, 5°C
Leave for 30mins in Fridge
Decant supernatant (IN SINK)



Use pipette P200, set at 110ul so that you aspirate ALL the sample! – don't froth!

Appendix L

DT Technical Data sheet for Naïve T cell panel

BD Biosciences

Technical Data Sheet (Eval)

Naïve T Cell Panel

Product Information **Custom Product**

Catalog Number: 626503
Size: 5 dried tubes per pouch
Volume: 1 test per tube

Description

CTT Dried tube: Dried 5-color cocktail in a 12 x 75 mm tube.

Format:	PerCP-Cy™5.5	PE	PE-Cy™7	APC	BV421
Specificity:	CD3	CD45RB	CD4	CD62L	CD45RA
Clone:	UCHT1	MT4	RPA-T4	DREG-56	HI100

Storage and Handling

Store tubes in their original foil pouch at 20°C-25°C†. Do not freeze the reagent or expose it to direct light at any time during storage or incubation with cells. Do not use reagents beyond the expiration date indicated on the label.

CAUTION: For pouches containing more than one tube, please take the following precautions.

- Due to the moisture sensitivity of the reagent, ensure the pouch is immediately and completely resealed after removing a tube. Do not remove the desiccant from the reagent pouch.
- Once the pouch is opened, use the remaining tubes within 1 month (when stored as directed).

†20°C-25°C is the "controlled room temperature" described as per USP34-NF29 and recommended by OSHA.

Usage

Refer to our website (www.bdbiosciences.com) or contact your local BD representative for the lyse/wash protocol for direct immunofluorescence.

1. After opening pouch, it is recommended to begin experiment immediately.
2. Staining blood samples with dried cocktails:
   a. Add 100 µL well-mixed whole blood* to the dried pellet at the bottom of the tube. Vortex gently to mix and incubate for 25 minutes in the dark at room temperature (20°C-25°C).
   b. After staining, proceed with sample processing (e.g. lysis, fixation, washing steps) using established laboratory procedures.
   c. Analyze samples on a flow cytometer, equipped with the appropriate filter and laser setup for antibody panel.

* To avoid serum interference when using reagents containing anti-kappa or anti-lambda antibodies, it is necessary to prewash the whole blood sample using at least 25 volumes of excess 1X PBS with 0.1% sodium azide (for example, 48 mL 1X PBS with sodium azide per 2 mL of whole blood to be washed). Mix well. Pellet cells by centrifugation and resuspend in 1X PBS with 0.1% sodium azide to the original volume.

Product Notices

- Cy™ is a trademark of GE Healthcare. This conjugated product is sold under license to the following patents: US Patent Nos. 5,268,486; 5,486,616; 5,569,587; 5,569,766; 5,627,027.

- For optimal and reproducible results, BD Horizon Brilliant™ Stain Buffer should be used anytime two or more BD Horizon Brilliant™ dyes are used in the same experiment. Fluorescent dye interactions may cause staining artifacts which may affect data interpretation. The BD Horizon Brilliant™ Stain Buffer was designed to minimize these interactions. More information can be found in the Technical Data Sheet of the BD Horizon Brilliant™ Stain Buffer (Cat. No. 563794/566349) or the BD Horizon Brilliant™ Stain Buffer Plus (Cat. No. 566385).

- BD Horizon Brilliant™ Ultraviolet 395, 496, 661, 737, 805, BD Horizon Brilliant™ Violet 421, 480, 510, 605, 650, 711, 750, 786, and BD Horizon Brilliant Blue 700 are covered by one or more of the following US patents: 8,110,673; 8,158,444; 8,227,187; 8,354,239; 8,362,193; 8,455,613; 8,575,303; 8,802,450; 8,841,072.

- BD Horizon Brilliant™ Stain Buffer is covered by one or more of the following US patents: 8,110,673; 8,158,444; 8,575,303; 8,354,239.

- Due to spectral differences between labeled cells and beads, using BD™ CompBeads can result in incorrect spillover values when used with BD Horizon BUV661, BUV737 or BV750 reagents. Therefore, the use of BD CompBeads or BD CompBeads Plus to determine spillover values for these reagents is not recommended. Different BUV661 or BUV737 reagents (eg, CD4 vs. CD45) can have slightly different fluorescence spillover therefore, it may also be necessary to use clone-specific compensation controls when using these reagents.

- Although every effort is made to minimize the lot-to-lot variation in the efficiency of the fluorochrome energy transfer, differences in the residual emission from BD Horizon™ BV421 may be observed. Therefore, we recommend that individual compensation controls be performed for every BD Horizon™ BV605 conjugate.

Manufactured by:
Becton, Dickinson and Company
BD Biosciences
11077 N Torrey Pines Rd.
La Jolla, CA 92037 USA

• PE-Cy7 conjugates: Although every effort is made to minimize the lot-to-lot variation in the efficiency of the fluorochrome energy transfer, differences in the residual emission from PE may be observed. Therefore, we recommend that individual compensation controls be performed for every PE-Cy7 conjugate.

• Source of all serum proteins is from USDA inspected abattoirs located in the United States.

• Caution: Reagents contain sodium azide. Sodium azide is harmful if swallowed. Keep away from food, drink, and animal feeding. Wear suitable protective clothing. If swallowed, seek medical advice immediately and show this container or label. Contact with acids liberates very toxic gas. Azide compounds should be flushed with large volumes of water during disposal to avoid deposits in lead or copper plumbing where explosive.

• Some reagents are bottled with ProClin® 950 (CAS No. 2682-20-4) and contains MIT (EC No. 220-239-6). Refer to safety data sheet for further information.

• Proclin is a registered trademark of Rohm and Haas Company.

Warning

May cause an allergic skin reaction.
Harmful to aquatic life with long lasting effect.

Wear protective gloves/eye protection. Wear protective clothing. Avoid breathing mist/vapours/spray. If skin irritation or rash occurs: Get medical advice/attention. IF ON SKIN: Wash with plenty of water. Dispose of contents/container in accordance with local/regional/national/international regulations.

Conditions

The information disclosed herein is not to be construed as a recommendation to use the product in violation of any patents. BD Biosciences will not be held responsible for patent infringement or the other violations that may occur with the use of our products. Purchase does not include or carry any right to resell or transfer this product either as a stand-alone product or as a component of another product. Any use of this product other than permitted use without the express written authorization of Becton Dickinson and Company is strictly prohibited. This product is not for use in therapeutic procedures and is not for resale.

For Research Use Only. This product is not for use in diagnostic or therapeutic procedures.

Appendix M

DT Technical Data Sheet for Treg panel

BD Biosciences

Technical Data Sheet (Eval)

Treg Panel

Product Information	**Custom Product**
Catalog Number:	626502
Size:	5 dried tubes per pouch
Volume:	1 test per tube

Description

CTT Dried tube: Dried 3-color cocktail in a 12 x 75 mm tube.

Format:	PE-Cy™7	Alexa Fluor® 647	BV421
Specificity:	CD4	CD127	CD25
Clone:	RPA-T4	HIL-7R-M21	2A3

Storage and Handling

Store tubes in their original foil pouch at 20°C-25°C†. Do not freeze the reagent or expose it to direct light at any time during storage or incubation with cells. Do not use reagents beyond the expiration date indicated on the label.

CAUTION: For pouches containing more than one tube, please take the following precautions.

- Due to the moisture sensitivity of the reagent, ensure the pouch is immediately and completely resealed after removing a tube. Do not remove the desiccant from the reagent pouch.
- Once the pouch is opened, use the remaining tubes within 1 month (when stored as directed).

†20°C-25°C is the "controlled room temperature" described as per USP34-NF29 and recommended by OSHA.

Usage

Refer to our website (www.bdbiosciences.com) or contact your local BD representative for the lyse/wash protocol for direct immunofluorescence.

1. After opening pouch, it is recommended to begin experiment immediately.
2. Staining blood samples with dried cocktails:
   a. Add 100 µL well-mixed whole blood* to the dried pellet at the bottom of the tube. Vortex gently to mix and incubate for 25 minutes in the dark at room temperature (20°C-25°C).
   b. After staining, proceed with sample processing (e.g. lysis, fixation, washing steps) using established laboratory procedures.
   c. Analyze samples on a flow cytometer, equipped with the appropriate filter and laser setup for antibody panel.

* To avoid serum interference when using reagents containing anti-kappa or anti-lambda antibodies, it is necessary to prewash the whole blood sample using at least 25 volumes of excess 1X PBS with 0.1% sodium azide (for example, 48 mL 1X PBS with sodium azide per 2 mL of whole blood to be washed). Mix well. Pellet cells by centrifugation and resuspend in 1X PBS with 0.1% sodium azide to the original volume.

Product Notices

- Cy™ is a trademark of GE Healthcare. This conjugated product is sold under license to the following patents: US Patent Nos. 5,268,486; 5,486,616; 5,569,587; 5,569,766; 5,627,027.

- For optimal and reproducible results, BD Horizon Brilliant™ Stain Buffer should be used anytime two or more BD Horizon Brilliant™ dyes are used in the same experiment. Fluorescent dye interactions may cause staining artifacts which may affect data interpretation. The BD Horizon Brilliant™ Stain Buffer was designed to minimize these interactions. More information can be found in the Technical Data Sheet of the BD Horizon Brilliant™ Stain Buffer (Cat. No. 563794/566349) or the BD Horizon Brilliant™ Stain Buffer Plus (Cat. No. 566385).

- BD Horizon Brilliant™ Ultraviolet 395, 496, 661, 737, 805, BD Horizon Brilliant™ Violet 421, 480, 510, 605, 650, 711, 750, 786, and BD Horizon Brilliant Blue 700 are covered by one or more of the following US patents: 8,110,673; 8,158,444; 8,227,187; 8,354,239; 8,362,193; 8,455,613; 8,575,303; 8,802,450; 8,841,072.

- BD Horizon Brilliant™ Stain Buffer is covered by one or more of the following US patents: 8,110,673; 8,158,444; 8,575,303; 8,354,239.

- Due to spectral differences between labeled cells and beads, using BD™ CompBeads can result in incorrect spillover values when used with BD Horizon BUV661, BUV737 or BV750 reagents. Therefore, the use of BD CompBeads or BD CompBeads Plus to determine spillover values for these reagents is not recommended. Different BUV661 or BUV737 reagents (eg, CD4 vs. CD45) can have slightly different fluorescence spillover therefore, it may also be necessary to use clone-specific compensation controls when using these reagents.

- Although every effort is made to minimize the lot-to-lot variation in the efficiency of the fluorochrome energy transfer, differences in the residual emission from BD Horizon™ BV421 may be observed. Therefore, we recommend that individual compensation controls be performed for every BD Horizon™ BV605 conjugate.

Manufactured by:
Becton, Dickinson and Company
BD Biosciences
11077 N Torrey Pines Rd,
La Jolla, CA 92037 USA

BD

S

• PE-Cy7 conjugates: Although every effort is made to minimize the lot-to-lot variation in the efficiency of the fluorochrome energy transfer, differences in the residual emission from PE may be observed. Therefore, we recommend that individual compensation controls be performed for every PE-Cy7 conjugate.

• Source of all serum proteins is from USDA inspected abattoirs located in the United States.

• Caution: Reagents contain sodium azide. Sodium azide is harmful if swallowed. Keep away from food, drink, and animal feeding. Wear suitable protective clothing. If swallowed, seek medical advice immediately and show this container or label. Contact with acids liberates very toxic gas. Azide compounds should be flushed with large volumes of water during disposal to avoid deposits in lead or copper plumbing where explosive.

• Some reagents are bottled with ProClin® 950 (CAS No. 2682-20-4) and contains MIT (EC No. 220-239-6). Refer to safety data sheet for further information.

• Proclin is a registered trademark of Rohm and Haas Company.

Warning

May cause an allergic skin reaction.
Harmful to aquatic life with long lasting effect.

Wear protective gloves/eye protection. Wear protective clothing. Avoid breathing mist/vapours/spray. If skin irritation or rash occurs: Get medical advice/attention. IF ON SKIN: Wash with plenty of water. Dispose of contents/container in accordance with local/regional/national/international regulations.

Conditions

The information disclosed herein is not to be construed as a recommendation to use the product in violation of any patents. BD Biosciences will not be held responsible for patent infringement or the other violations that may occur with the use of our products. Purchase does not include or carry any right to resell or transfer this product either as a stand-alone product or as a component of another product. Any use of this product other than permitted use without the express written authorization of Becton Dickinson and Company is strictly prohibited. This product is not for use in therapeutic procedures and is not for resale.

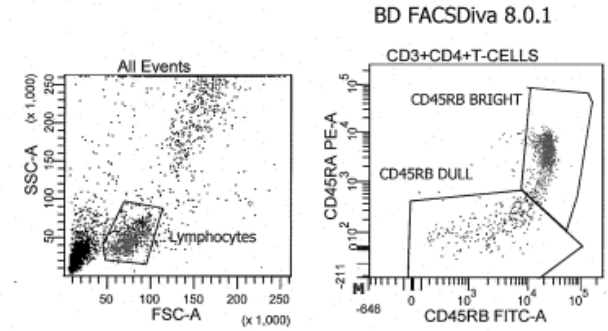
For Research Use Only. This product is not for use in diagnostic or therapeutic procedures.

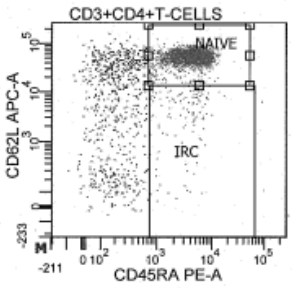
Manufactured by:
Becton, Dickinson and Company
BD Biosciences
11077 N Torrey Pines Rd.
La Jolla, CA 92037 USA

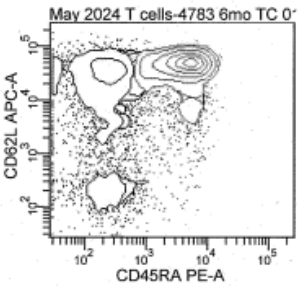
BD

Appendix N

NHS gating stratergies for T-cells and Treg panels

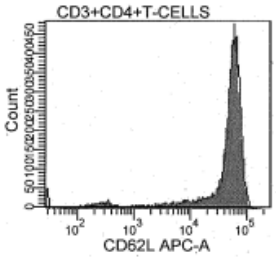






Tube: [redacted]

Population	#Events	%Parent	%Total
All Events	196,720	####	100.0
Lymphocytes	18,164	9.2	9.2
CD3+CD4+T-CELLS	10,220	56.3	5.2
CD45RB BRIGHT	7,532	73.7	3.8
CD45RB DULL	2,475	24.2	1.3
NAIVE	6,856	67.1	3.5
IRC	505	4.9	0.3



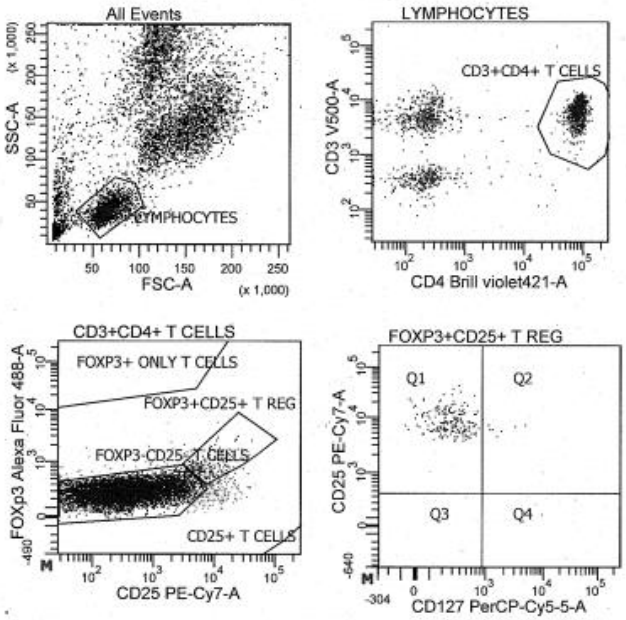
Experiment Name: T Cells April 2024_001
Specimen Name: May 2024 T cells
Tube Name: [redacted]
Record Date: May 1, 2024 11:34:12 AM

Population	#Events	%Parent	CD62L APC-A Median
All Events	196,720	####	85
Lymphocytes	18,164	9.2	39,575
CD3+CD4+T-CELLS	10,220	56.3	49,363
CD45RB BRIGHT	7,532	73.7	53,257
CD45RB DULL	2,475	24.2	14,129
NAIVE	6,856	67.1	54,883
IRC	505	4.9	4,445

T-Cell Panel

Printed on: Wed May 15, 2024 02:22:00 BST

BD FACSDiva 8.0.1



Tube: [redacted]

Population	#Events	%Parent	%Total
All Events	73,702	####	100.0
LYMPHOCYTES	17,847	24.2	24.2
CD3+CD4+ T CELLS	10,229	57.3	13.9
FOXP3+ ONLY T CE	0	0.0	0.0
FOXP3+CD25+ T R	200	2.0	0.3
CD4+CD127- T	198	99.0	0.3
CD62L+ T-REG	0	0.0	0.0
CD62L- T-REG	121	60.5	0.2
CD45RA- T-REC	121	60.5	0.2
CD45RA+ T-RE	0	0.0	0.0
HLA-DR- T-REG	197	98.5	0.3
HLA-DR+ T-REC	0	0.0	0.0
Q1	193	96.5	0.3
Q2	7	3.5	0.0
Q3	0	0.0	0.0
Q4	0	0.0	0.0
CD25+ T CELLS	0	0.0	0.0
FOXP3-CD25- T CE	9,728	95.1	13.2
CD62L+ T-CELL	0	0.0	0.0
CD62L- T-CELL	9,607	98.8	13.0
CD45RA- T-CEL	9,721	99.9	13.2
CD45RA+ T-CEI	3	0.0	0.0
HLA-DR- T-CEL	8,768	90.1	11.9
HLA-DR+ T-CEL	0	0.0	0.0

T-REG PANEL

Printed on: Wed May 15, 2024 02:20:02 BST