# Exploration of Cellular Signaling Patterns for the Stratification of Patients with Rheumatoid Arthritis

Lucius Bader

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2020

UNIVERSITY OF BERGEN

# Exploration of Cellular Signaling Patterns for the Stratification of Patients with Rheumatoid Arthritis

Lucius Bader


Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 16.09.2020

Year: 2020

Title: Exploration of Cellular Signaling Patterns for the Stratification of Patients with Rheumatoid Arthritis

Name: Lucius Bader

Print: Skipnes Kommunikasjon / University of Bergen

# Scientific environment

Haukeland University Hospital, Bergen, Norway:

- The Bergen group of Epidemiology and Biomarkers in Rheumatic Disease, BEaBiRD, at the Department of Rheumatology
- Department of Neurology
- The Neuroimmunology Laboratory

University of Bergen, Norway:

- The Faculty of Medicine and Dentistry, Department of Clinical Science
- Centre for Cancer Biomarkers, CCBIO
- The Flow and Mass Cytometry Core Facility
- The Broegelmann Research Laboratories and Research School of Inflammation
- The Faculty of Mathematics and Natural Sciences, Department of Mathematics

# Acknowledgements

Thank you.

**Supervisors.** *Christian*, you must have gotten some grey hair on this (the candidate did!). Or lost some. Emptied university accounts. No more space in liquid nitrogen. An eternal candidate. You have been a patient and forgiving man. *Clara*, you are the everything-is-possible, making-everything-possible woman. A facilitator from A to Å. You disappeared but you are still just around the corner. *Sonia*, you are a true master and a friend. The candidate admires your handling of the (researcher) world's craziness.

**Department of Rheumatology and BEaBiRD.** *Patients*, it's all about you, isn't it? And you are willing to bleed even a bit more. *Tove, Signe Lise, Torill, nurses at the outpatient clinic – and all other employees*, you run until you drop, and you do it with a smile. *Colleagues*, you listen to wild stories about light bulbs and about how to be a bird. And show true empathy through all the candidate's states, from enthusiasm to desperation. *Helse Vest*, the candidate doesn't know you personally (maybe you *just are* you and me), but still you grant a PhD scholarship and project money. *Miriam*, you have suffered and worked hard. I really hope you'll be rewarded.

**Neuroimmunology laboratory.** *Hanne Linda, Mette, Liesbeth, Kibret*, you guide the ignorant through the world of RPMI, DMSO, PFA, pipetting and nanogramming, and won't get more than a pack of Lebkuchen and 3 lines in some acknowledgements. You deserve better (well, at least it is 4 lines now).

**Core facility for flow cytometry**. *Brit*, noble lady over machines and advent calendars, you are the heroine of all cytometry candidates.

**CCBIO and fellow PhD-students.** *Stein-Erik and Jørn*, you know all-the-things, you know how all-the-things work, you know how all-the-things are done. And you share. *Gerd,* running on eggs and bacon, you really get done things. Don't eat your sheep.

*Bjørn-Tore*, doubts must have crossed your mind (so many things seem to cross your mind!). You didn't have to be, but you have always been a supporter.

**Broegelmann.** *Marianne, Silke, Roland, Karl*, sometimes we speak in different tongues, but ultimately, we share the same language. You endured the stubborn candidate. You were there with salary and project funds, with a helping hand and reflection.

**Mathematics, bioinformatics.** *Nello and Morten* (*or NM2B)*, you guys are amazing. The candidate may share some fundamental brain functions with you (eat, move, sleep), but all the rest is just so different. Luckily.

**Family and friends and woods and music.** *CoRE, Legenes Julekor and 454.0,* you are my music. *Moving outside*, into the woods – the escape to real life – which also *Adnan and Øyvind* with families are a big part of. *MMPP*, you always let me – be. *Andrej, Charlie und Mati*, you are the UP and the DOWN, the LEFT and the RIGHT, the FRONT and the BACK – usually all at the same time. And if people and things and thoughts are spread throughout the universe, you are the middle, the gravitational center, the powerful force that keeps people and things and thoughts together, *Monika.*

Thank you.

The candidate, 13.03.20.

# Technical terms and abbreviations

**Technical terms**

| | |
|---|---|
| Apoptosis | "Ordered" programmed cell death in multicellular organisms |
| Necroptosis | "Messy" or inflammatory programmed cell death |
| Arcsinh | A hyperbolic arcsine function for the transformation of cytometry data, which allows for the transformation of negative raw values, and for linear-type data display for values around zero within the so-called scale argument |

**Abbreviations**

| | |
|---|---|
| ACPA | Anti-citrullinated peptide/protein antibodies |
| ACR | American College of Rheumatology |
| ADAb or ADA | Anti-drug antibody |
| ADL | Adalimumab, a therapeutic human antibody against TNF |
| Akt | Protein kinase B |
| BAb | Binding antibody |
| Bc | B cell(s), express CD19 and CD20 |
| BM | Bone marrow |
| bDMARD | biologic disease-modifying antirheumatic drug(s) |
| CD | Cluster of differentiation |
| CD120a, CD120b | Tumor necrosis factor receptors 1 and 2, syn. TNFR1, TNFR2 |
| CITRUS | Cluster identification, characterization, and regression algorithm |
| cM | classical monocyte(s), express CD14 |
| CPT | Cell preparation tube, BD Vacutainer® CPT™, containing Na-citrate, no heparin |
| CRD | Cysteine-rich domain |
| CRP | C-reactive protein, a marker of inflammation |
| CyTOF | Cytometry by time-of-flight, mass cytometry |

| | |
|---|---|
| CZP | Certolizumab pegol, a therapeutic PEGylated fab-fragment against TNF |
| Dc | Dendritic cell(s), express CD123 and/or CD11c and/or HLA-DR |
| DD | Death Domain |
| DMARD | Disease-modifying antirheumatic drug(s) |
| DNA | Deoxyribonucleic acid |
| ELISA | Enzyme-linked immunosorbent assay |
| ESR | Erythrocyte sedimentation rate, a marker of inflammation |
| ETN | Etanercept, a recombinant therapeutic receptor binding TNF |
| EULAR | European League Against Rheumatism |
| GC | Glucocorticosteroids, e.g. prednisolone |
| GLM | Golimumab, a therapeutic human antibody against TNF |
| GM-CSF | Granulocyte-macrophage colony-stimulating factor |
| INFα | Interferon alpha |
| IFX | Infliximab, a therapeutic chimeric antibody against TNF |
| JIA | Juvenile idiopathic arthritis |
| kDa | kilo-Dalton (unit for molecular mass) |
| LPS | Lipopolysaccharide |
| LT | Lymphotoxin, syn. tumor necrosis factor beta (TNF-ß) |
| M | Monocyte(s), express CD14 and/or CD16 and/or CD11c |
| Mab, -mab | monoclonal antibody |
| mDc | myeloid dendritic cell(s), express CD11c and HLA-DR |
| MS | Multiple sclerosis or mass spectrometry |
| MTX | Methotrexate |
| Nab | Neutralizing antibody |
| NKc | Natural killer cell(s), express CD56 |
| NK Tc | Natural killer T cell(s), express CD3, CD8, CD16 |
| OMIP | Optimized multicolor immunophenotyping panel |
| PBL | Peripheral blood leukocytes (including granulocyte populations) |
| PBMC | Peripheral blood mononuclear cells (excluding granulocyte populations) |

| | |
|---|---|
| pDc | plasmacytoid dendritic cell(s), express CD123 and HLA-DR |
| PFA | Paraformaldehyde, a fixative |
| RA | Rheumatoid arthritis |
| RF | Rheumatoid factor |
| RGA | Reporter-gene assay |
| sDMARD | synthetic disease-modifying antirheumatic drug(s) |
| SMD | Small molecule drug(s) |
| SOP(s) | Standard operating procedure(s) |
| SPADE | Spanning-tree Progression Analysis of Density-normalized events |
| sTNF | soluble tumor necrosis factor (trimer) |
| TACE | tumor necrosis factor (alpha) converting enzyme |
| Tc | T cell(s), T lymphocyte(s), express CD3 and CD4 or CD8 |
| tDMARD | targeted disease modifying anti-rheumatic drug |
| tmTNF | transmembrane tumor necrosis factor (trimer) |
| TNF | Tumor necrosis factor, syn. tumor necrosis factor alpha (TNFα) |
| TNFi | Tumor necrosis factor inhibitors |
| TNFR1, TNFR2 | Tumor necrosis factor receptors 1 and 2, syn. CD120a, CD120b |
| TNFRSF | Tumor necrosis factor receptor superfamily |
| TNFSF | Tumor necrosis factor superfamily |
| TOF | Time-of-flight, see CyTOF |
| TRAF | TNF receptor associated factors |
| tSNE | t-stochastic neighbor embedding |
| viSNE | Visualization of t-stochastic neighbor embedding |

# List of publications

1. Bader LI, Solberg SM, Kaada SH, Bolstad N, Warren DJ, Gavasso S, Gjesdal CG, Vedeler C. Assays for infliximab drug levels and antibodies: a matter of scales and categories. Scand J Immunol. 2017.

2. Gullaksen SE, Bader L, Hellesøy M, Sulen A, Fagerholt OEE, Engen CB, Skavland J, Gjertsen BT, Gavasso S. Titrating complex mass cytometry panels. Cytometry Part A. 2019.

3. Bader L, Gullaksen SE, Blaser N, Brun M, Sulen A, Vedeler C, Gram Gjesdal C, Gavasso S. Candidate markers for stratification and classification in rheumatoid arthritis. Frontiers in Immunology. 2019.

# Abstract

*Introduction* Rheumatoid arthritis (RA) is a chronic inflammatory disease, characterized by synovial inflammation that, if not treated early and efficiently, causes joint damage. The proinflammatory cytokine tumor necrosis factor (TNF) plays a central role in the pathogenesis of RA and is the target of treatment with TNF inhibitors. TNF inhibitors are generally effective and well-tolerated. However, up to one-third of patients are primary non-responders, and responses in up to one-third of initial responders abate over time. Currently, there are no predictive biomarkers for treatment with TNF inhibitors. TNF inhibitor drug levels and anti-drug antibodies (ADAb) are indicative of secondary treatment responses, but these markers are not standardized for clinical application. Previous studies have indicated the potential of single-cell profiling by flow or mass cytometry for patient stratification in RA and in other autoimmune conditions. Distinct signaling patterns have been found in leukocytes of RA patients before and during treatment with TNF inhibitors in exploratory and proof-of-principle studies.

The aim of this thesis was to explore and compare existing markers for TNF inhibitor drug responses, to set up a methodological background for mass cytometry experiments and finally to explore signaling signatures in immune cell subsets from RA patients compared to healthy individuals, with an emphasis on TNF signaling.

*Material, methods and results* First, we explored existing assays for TNF inhibitor drug levels and for ADAb in sera from 107 patients with inflammatory diseases treated with the TNF inhibitor infliximab. We found that the assays measured on different scales and that the agreement between quantitative results was limited. However, inter-assay differences could partially be overcome by assay-individualized translations of quantities into categories, which is also necessary for meaningful clinical application **(paper 1)**. Second, we established a basis for mass cytometry experiments, including the extensive collection of biobank material and patient data.

Methodological work in the design and titration of antibody panels for mass cytometry was carried out to provide a hierarchical titration method for complex mass cytometry panels, which takes in account abundancies, sources of signal spillover and non-specific antibody binding **(paper 2)**. Last, we explored signaling patterns in heterogeneous immune cells from 20 newly diagnosed RA patients and from 20 healthy donors, with a focus on TNF signaling. In an automated data analysis pipeline, 18 of 20 RA patients and 17 of 20 healthy donors were correctly classified based on their signaling patterns **(paper 3)**.

*Conclusion* RA is a heterogeneous disease with a plethora of treatment options, and patients might profit from more exact classification and stratification. This thesis highlights the lack of classification and stratification markers, and shows, how single cell profiling by mass cytometry may contribute to the search for such markers. Methodological aspects such as antibody panel design and approaches for the analysis of high-dimensional data are emphasized. The core results of the thesis show that newly diagnosed RA patients can be classified correctly with relatively high precision based on signaling patterns in single cells, when compared to healthy donors.

The mass cytometry platform adds many dimensions to “cytomics” of heterogenous cell suspensions and tissues. While studies on malignancies as well as physiology and development of the immune system dominate the field, rheumatic diseases are currently relatively underrepresented. The door for further and deeper study of rheumatic diseases and signaling far beyond the TNF pathway is wide open.

# Introduction and methods

## 1. Rheumatoid arthritis and tumor necrosis factor

Rheumatoid arthritis (RA) is a chronic autoimmune, inflammatory disease. Synovitis in small- and medium-sized joints is the hallmark of RA, causing the typical symptoms of inflammation in affected areas – *calor, rubor, dolor, tumor, functio laesa* – and, if not treated early and efficiently, joint damage and destruction. RA affects women more frequently than men (3:1), with a peak prevalence at around 65 years of age and a total prevalence of about 0.5-1% in developed countries [1]. RA is considered to be a multifactorial disease with genetic/hereditary, environmental and infectious risk factors.

| **ACR/EULAR 2010 criteria for rheumatoid arthritis** | |
|---|---|
| Joint involvement | (0-5) |
| 1. One medium-to-large joint | (0) |
| 2. Two to ten medium-to-large joints | (1) |
| 3. One to three small joints (large joints not counted) | (3) |
| 4. Four to ten small joints (large joints not counted) | (4) |
| 5. More than ten small joints (at least 1 small joint) | (5) |
| Serology | (0-3) |
| 1. Negative RF and negative ACPA | (0) |
| 2. Low positive RF or low positive ACPA | (2) |
| 3. High positive RF or high positive ACPA | (3) |
| Acute-phase reactants | (0-1) |
| 1. Normal CRP and normal ESR | (0) |
| 2. Abnormal CRP or abnormal ESR | (1) |
| Duration of symptoms | (0-1) |
| 1. Less than six weeks | (0) |
| 2. Six weeks or more | (1) |
| Points are shown in parenthesis. Cut-point for rheumatoid arthritis six points or more. Patients can also be classified as having rheumatoid arthritis if they have (a) typical erosions, (b) long-standing disease previously satisfying the classification criteria. | |

**Table 1: ACR/EULAR 2010 criteria for rheumatoid arthritis.**
RF=Rheumatoid Factor
ACPA=Anti-Citrullinated Peptide/Protein Antibodies
CRP=C-Reactive Protein
ESR=Erythrocyte Sedimentation Rate.

The diagnosis “rheumatoid arthritis” describes a collection of disease phenotypes rather than a single specific disease entity. This is reflected by the American College of Rheumatology/European League Against Rheumatism 2010 criteria for RA **(Table 1)**, which point to the “typical” RA patient with multiple inflamed small joints, positive rheumatoid factor (RF) and/or autoantibodies against citrullinated proteins (ACPA), elevated acute-phase reactants and a certain duration of symptoms. On the other hand these criteria allow for the same diagnosis in a rather wide range of patients, e.g. for patients without small-joint affection, without RF/ACPA or without elevated acute-phase reactants [2].

A curing treatment for RA does not exist. There are, however, promising attempts at more causal therapeutic approaches, e.g. through targeting dendritic cells in order to induce long-lasting immune tolerance in patients [3]. “Treat to target” is the major treatment principle for RA, in which the most common treatment target would be either remission or low disease activity. Empiric RA treatment usually starts with glucocorticosteroids (GC) combined with one or several synthetic disease-modifying antirheumatic drugs (sDMARD), e.g. methotrexate. Second-line treatments in case of adverse events or lack of therapeutic response are biologic disease-modifying antirheumatic drugs (bDMARD) and, more recently, small-molecule drugs (SMD), also referred to as targeted DMARD (tDMARD) [4].

One of the most central players in the pathogenesis of RA – and of several other autoimmune, inflammatory diseases – is the pro-inflammatory cytokine “tumor necrosis factor” (TNF).

## 1.1. A brief history of tumor necrosis factor

Since the end of the 19th century, cancer scientists explored the phenomenon of “hemorrhagic necrosis” – the ability of bacterial endotoxins to induce regression of malignant tumors [5]. At that time, one of several synonyms for bacterial endotoxins was “tumor-necrotizing toxin”, reflecting the early assumption that bacterial

endotoxins featured direct tumor-necrotizing abilities [6]. This assumption was shown to be wrong when Carswell et al. in 1975 demonstrated that endotoxin-treated mice developed a serum factor, which induced tumor necrosis after injection into sarcoma-transplanted, endotoxin-free mice. They called this factor the "tumor necrosis factor" [7]. Direct evidence for a macrophage-derived cytotoxic cytokine and details about its biochemical structure were provided by Aggarwal et al. in 1985. They called this cytokine TNF *alpha* to separate it from Carswell's more undefined "tumor necrosis factor" and to distinguish it from its sibling-cytokine lymphotoxin (LT), then named TNF-beta. It has later been suggested that the terms TNF and LT should be used for the sake of simplicity and to avoid confusion [8]. The work of Aggarwal et al. was paralleled by research of Beutler et. al in 1985, who reported on a cytokine called cachectin, which subsequently was shown to be identical with TNF [9-11]. The systemic effects attributed to cachectin alias TNF – fever, shock and catabolism/cachexia – made it clear that the cytokine could not be used as a universal systemic therapeutic for all kinds of cancer tumors [12]. Shortly after the discovery of TNF/cachectin and LT, their genes were cloned, which revealed that TNF and LT were part of one gene family, the TNF superfamily [13, 14].

## 1.2. TNF, its receptors and their superfamilies

TNF is produced predominantly by activated macrophages, lymphocytes, natural killer cells and granulocyte populations, and – upon certain stimuli – also by several tissue cell types. TNF is a protein with a molecular mass of about 17kDa in its monomeric form and an isoelectric point of 5.3. It contains one disulfide bridge based on two cysteines [9]. In vivo, TNF is expressed as a transmembrane type II protein trimer (tmTNF), which can be cleaved as soluble TNF (sTNF) by the help of TNF alpha converting enzyme (TACE or ADAM17) [15]. The differentiation between tmTNF and sTNF has several important implications for the function and regulation of TNF. First of all, it allows for TNF signaling through direct cell-to-cell contact as well as remote effects by circulating sTNF. Secondly, tmTNF has been shown to be

capable of reverse signaling. In reverse signaling, the ligand – tmTNF – functions as a receptor itself, altering the state of the cell it is anchored to when binding a cognate TNF receptor on a target cell [16]. Thirdly, only tmTNF is able to fully activate TNF receptor 2, one of the two TNF receptors [17]. TNF is a member of the TNF superfamily (TNFSF), a group of 19 ligands sharing homologies in DNA sequences and functionality regarding their role in both cell survival and cell death [18, 19]. Virtually all ligands of the TNF superfamily share pro-inflammatory features, often based on similar signaling pathways.

The cognate counterpart to the TNF superfamily ligands is the TNF receptor superfamily (TNFRSF), consisting of 29 receptors. Receptors from the TNF receptor superfamily are type I transmembrane proteins and contain typically a varying number of cysteine-rich domains (CRD) extracellularly. Of intracellular features, the death domains (DD) and TNF receptor associated factors (TRAF) should be mentioned. TNF itself has two cognate receptors, TNF receptor 1 and TNF receptor 2 (TNFR1, TNFR2 or, following the cluster of differentiation nomenclature, CD120a and CD120b). While TNFR1 is present on almost all human cells, TNFR2 is predominantly expressed on immune cells [20]. “Ligand passing” between the two different TNF receptors, possibly through the formation of TNF receptor heterocomplexes, may contribute to the regulation of TNF responsiveness of a cell [21].

## 1.3. TNF signaling

The most characteristic effects of TNF signaling have been described as “live or let die” [22]. Depending on setting and cell type, TNF can prompt a more pro-inflammatory behavior with increased proliferation, differentiation and survival in the target cell or induce apoptosis/necroptosis **(Figure 1)**. TNFR1 – as opposed to TNFR2 – contains an intracellular death domain, which can induce cell death through caspases. The pro-apoptotic properties of TNFR2 are weaker due to the lack of death domains. Apoptosis is, however, not the most common outcome for TNF signaling

through neither TNFR1 nor TNFR2. Both receptors are strong activators of the canonical NFkB (nuclear factor kappa-light-chain-enhancer of activated B cells) pathway. Upon dual phosphorylation and degradation of the IkBa (inhibitor of kappa-light-chain-enhancer in B cells, alpha), NFkB is translocated into the nucleus of the cell and induces NFkB-sensitive genes, which – depending on cell type, status and cofactors – results in cell proliferation and differentiation with subsequent stimulatory effects on immune activation and inflammation. Gene products from canonical NFkB signaling also prevent the initiation of apoptosis by inhibiting the prolonged activation of Jun-N-terminal kinases (JNK) and caspases [23]. TNF furthermore induces mitogen-activated protein kinases (MAPK), JNK and phosphatidylinositol 3-kinases (PI3K), all of which can interact with the NFkB pathway and all of which, under given circumstances, can protect the cell from pro-apoptotic effects of TNF [19, 20, 24].

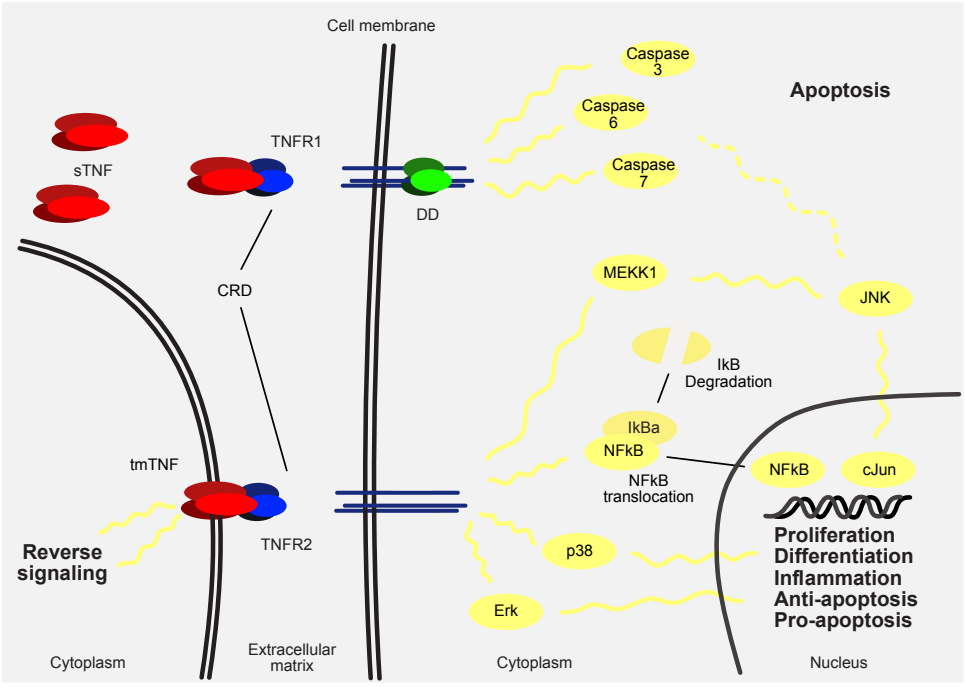



**Figure 1: TNF signaling.** Soluble TNF (sTNF) and transmembranous TNF (tmTNF) binding the TNF receptors (TNFR1+2), which contain cysteine-rich domains (CRD). TNFR1 contains also an intracellular death domain (DD). Yellow: a choice of signaling pathways and epitopes which are relevant for this thesis.

# 2. Tumor necrosis factor inhibitors (TNFi)

Due to its pivotal role in inflammatory diseases, TNF has early been recognized as a target for therapeutic intervention [25-28]. One of the first therapeutic antibodies following Köhler and Milstein's groundbreaking work in the production of monoclonal antibodies, was the chimeric anti-TNF-antibody "cA2", called infliximab after commercialization [29, 30]. Of today, millions of patients have received and do receive treatment with TNF inhibitors (TNFi): infliximab and its biosimilar siblings, the fully human antibodies adalimumab with biosimilars and golimumab, the PEGylated humanized antibody (Fab'-) fragment certolizumab pegol and the recombinant TNF receptor p75 fusion protein etanercept with biosimilars **(Figure 2)**. The widespread use of TNF inhibitors in the treatment of inflammatory conditions has resulted in a need to address three tightly connected issues regarding this group of drugs: *costs*, *side effects* and *treatment responses*:

*Costs.* TNF inhibitors are costly biopharmaceuticals. After their introduction in Norway in the end of the 1990s, expenses for TNF inhibitors have grown almost every year to about 2.6 billion NOK/317 million USD in 2016 [31]. Similar developments have been seen in other Western countries, setting up societal resources against the needs of the individual. Pharmacoeconomic considerations and the complex issue of cost effectiveness of TNF inhibitors have been addressed in several studies [32-34].

*Adverse events.* Side effects and adverse events are not uncommon in the treatment with TNF inhibitors and can be caused by the drug directly (e.g. hypersensitivity reactions) or indirectly by drug effects (e.g. infections). The term "biotoxicity" describes the spectrum of unwanted reactions to biopharmaceuticals, ranging from immunostimulation over immunogenicity, immune deviation and cross reactivity to non-immunological based reactions [35]. Hence, screening for pre-existing infectious conditions and tight monitoring for biotoxicity have become part of daily clinical routine and are of continued research interest [36-39].

*Treatment responses.* Because of the issues of cost and adverse events, the efficient evaluation of TNFi treatment responses has become more and more important. Clinical tools such as disease activity scores combined with markers of inflammation (C-reactive protein and erythrocyte sedimentation rate) are applied in daily clinical practice, and tremendous efforts have been made to identify and establish more specific and also predictive (bio)markers of TNF inhibitor effects [40-43].

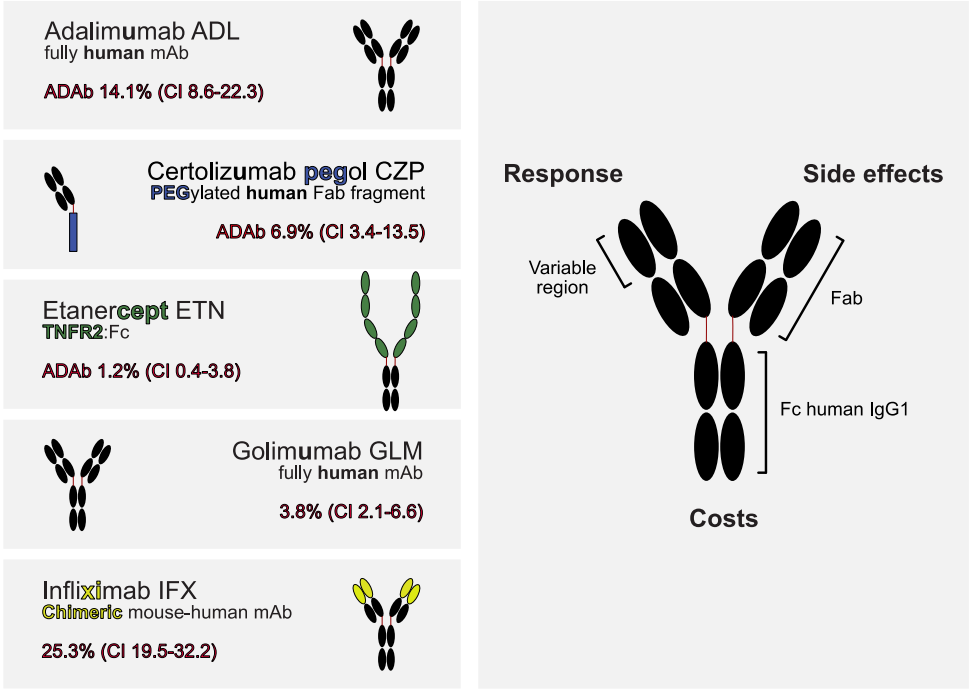



**Figure 2: TNF inhibitors.** The five TNF inhibitors, their structure and composition (left). Prevalence of anti-drug antibodies (ADAb) in percent (red), based on meta-analytic data [44]. The antibodies adalimumab, golimumab and infliximab are IgG1 antibodies with a strong complement activating capability and high affinity to Fc receptors on phagocytic cells. Composition of an IgG1 antibody, and the three challenges connected to the use of TNF inhibitors (right).

# 3. Response and non-response to tumor necrosis factor inhibitors

For the majority of RA patients treated with a TNF inhibitor, the treatment contributes to a significant reduction of inflammation and long-term damage, and to an improvement in quality of life [45]. However, a significant number of patients experience insufficient therapeutic responses. Up to one third of RA patients are primary non-responders – they do *not respond to TNF inhibitors at all* [2, 46]. While primary non-response in some patients may simply be caused by insufficient drug levels, it has also been suggested that disease mechanisms in a subgroup of patients are linked to TNF signaling to a lesser degree, resulting in reduced TNF inhibitor responses [47]. Of patients who primarily respond to treatment, another third experience *abating responses over time* – a so-called secondary loss of response or secondary non-response. Most cases of secondary non-response are attributed to immunogenic features of the biologic drug with subsequent neutralization of the drug through the host's immune system. Although drug immunogenicity is the only identified cause of secondary non-response, not all cases are attributable to the immunogenic features of the drug [48].

## 3.1. Drug immunogenicity

Drug immunogenicity as a cause of secondary non-response and, not least, of side effects and adverse events, is a well-known problem of biopharmaceuticals such as monoclonal antibodies, recombinant proteins, cytokines and hormones [49]. Throughout the history of treatment with biopharmaceuticals, there have been many examples of immunogenicity-caused loss of response and adverse events. Prominent examples are total red cell aplasia under treatment with recombinant erythropoietin due to immunogenic reactions against both supplied and self-produced erythropoietin in patients with chronic renal failure and renal anemia, or relapses in patients with multiple sclerosis due to the development of anti-drug antibodies against the therapeutic cytokine interferon-ß [49-51]. For TNF inhibitors, immunogenicity was

suspected and identified as a cause of secondary loss of response already in early clinical trials on the first TNF inhibitor, cA2/infliximab [52, 53].

Many factors contribute to or influence the immunogenic potential of a drug. These can be divided into drug-related and patient-related factors. Drug-related factors range from manufacturing processes, biochemical composition and size to formulation of the drug, while patient-related factors include patient age, gender, weight and genetics, individual drug dose and way of admission as well as type of disease, disease states and co-medication [49, 54-56]. Immunogenicity is a type ß immunotoxic reaction, also characterized as a delayed hypersensitivity reaction to the drug. T cell-dependent and T cell-independent pathways have been suggested as causes of immunogenicity. In both pathways, the biologic drug is detected by antigen-presenting cells, such as dendritic cells or naïve B cells. B cells can develop directly into plasma cells excreting specific IgM and, to a lesser degree IgG (T cell-independent). In T cell-dependent cases, antigen-presenting cells will activate T cells, which then prime B cells to develop into plasma cells with predominant production of IgG anti-drug antibodies (ADAb) [49, 54].

In a meta-analysis including over 14.000 patients, the overall prevalence of ADAb in patients treated with a TNF inhibitor was 12.7% and reduced the odds for treatment response by 67%, while at the same time increasing the odds (OR 3.25) for injection/infusion-related adverse events [44]. The prevalence of ADAb separately specified for each TNF inhibitor is given in figure 2.

Several primary and meta-analytic studies have shown that concomitant use of synthetic disease-modifying antirheumatic drugs (sDMARD), such as e.g. methotrexate together with a TNF inhibitor, reduces the odds for ADAb formation and immunogenic reactions significantly [44, 57, 58].

## 3.2. Current biomarkers for TNF inhibitor treatment responses

Up to one-third of RA-patients are primary non-responders to TNF inhibitors. *Predictive* (bio)markers for TNF inhibitor responses could spare these patients for the time spent on try-and-fail and potential adverse events, as well as the society for unnecessary drug expenses. The spectrum of predictive candidate (bio)markers includes early radiographic findings, smoking exposure, markers of inflammation, expression of autoantibodies and a long list of different genotypes. Radiographic findings and autoantibodies such as RF/ACPA are routinely applied mostly for diagnostic purposes and to make more general statements about expected disease severity and needed treatment intensity. Autoantibodies may have a certain relevance as biomarkers for the choice of B-cell depleting therapy (rituximab) and the CTLA-4 co-stimulation inhibitor abatacept [59]. Yet, a recent systematic review on biomarkers concluded that none of the studied predictors added value to clinical decision-making regarding treatment with TNF inhibitors or other biologic drugs [40].

On the other hand, *monitoring* (bio)markers for TNF inhibitor responses after treatment initiation do exist and are currently applied in clinical practice: drug levels and anti-drug antibodies (ADAb).

### *3.2.1. Drug levels*

TNF inhibitor serum levels reflect the pharmacokinetics of the drug and are again influenced by drug-related factors (way of admission, frequency, dose, formulation, drug storage) and patient-related factors (gender, weight, comorbidity, comedication). Many different assays – commercial and non-commercial – are available to measure drug levels, for the most immunoassays such as enzyme immunoassays (EIA), enzyme-linked immunosorbent assays (ELISA) and radio-immunoassays (RIA) [47]. A more functional approach is chosen in reporter gene assays (RGA), where levels of functionally available drug are measured by the help of a TNF-sensitive reporter gene cell line [60].

An important issue about drug levels is the time point of measurement, especially for TNF inhibitors with long dosing intervals. For drugs administered with intervals over

2 weeks, drug levels should therefore be measured as “trough” levels, at the very end of the administration interval, right before the next administration.

### *3.2.2. Anti-drug antibodies (ADAb)*

Similar modalities exist for the detection and quantification of ADAb, both immunoassays and reporter-gene assays **(Figure 3)**. However, quantities of ADAb are not measured in standardized units, and direct inter-assay comparisons of ADAb-titers are therefore not possible. Two important characteristics have to be considered for the understanding and interpretation of ADAb assays: drug-tolerance and assay functionality.

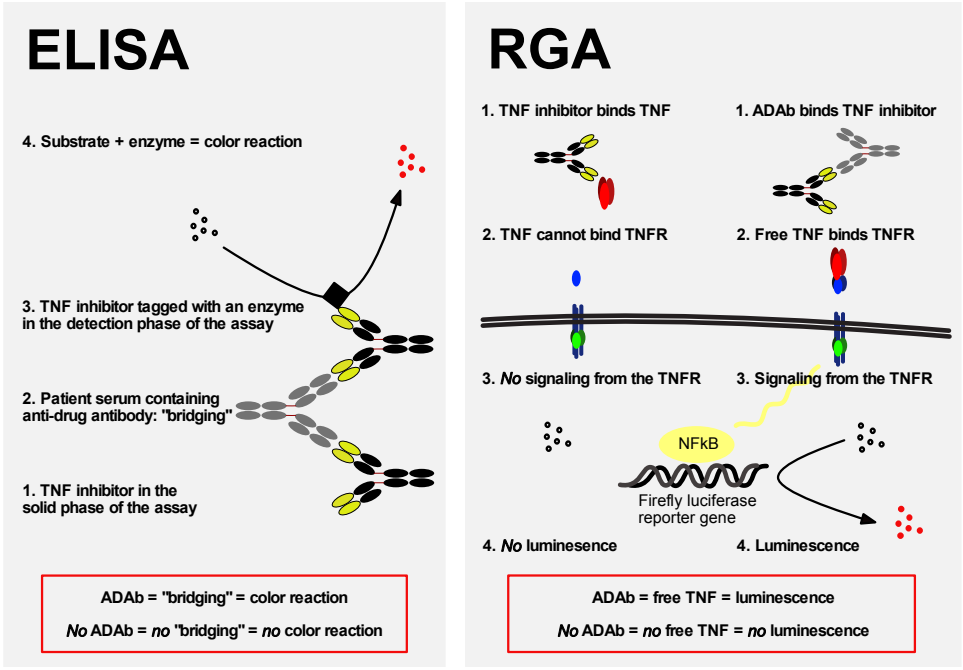



**Figure 3: Bridging ELISA and RGA for the detection of ADAb.** Figures adapted from Bendtzen [47].

*Drug-tolerance:* in the presence of the drug, ADAb will be bound to the drug and therefore “invisible” for many detection assays. Only excess ADAb after saturation of active drug will be detected by these “drug-sensitive” assays. Consequently, in a clinical setting ADAb assays are usually only carried out, if drug levels are very low. The problem of drug-ADAb interference can partially be overcome by dissociation

steps, where drug and ADAb are separated e.g. by acid dissociation. The clinical relevance of very low levels of drug antibodies is uncertain, and high-sensitive ADAb-detection to this point mostly of scientific interest [61]. For clinical purposes the primary question will be whether drug levels are sufficient, and ADAb detection will only be relevant in the case of low drug levels [62].

*Assay functionality:* not all ADAb necessarily hamper the functioning of the drug. One can therefore differentiate ADAb further into *binding* and *neutralizing* antibodies. Binding antibodies (BAb) bind the drug either at functionally unimportant sites or with low affinity, while neutralizing antibodies (NAb) bind the drug in ways which functionally neutralize the drug. Certain assays, e.g. solid-phase assays, will therefore detect ADAb with high sensitivity but low specificity for neutralizing abilities, while functional assays, such as reporter-gene assays, specifically detect neutralizing antibodies [60, 63, 64].

### *3.2.3. Challenges to drug levels and ADAb as clinical biomarkers*

Drug levels are quantified in standardized units, allowing for direct inter-assay comparisons of test results. However, despite good correlations between different assays, the inter-assay agreement (= true absolute accordance) has been shown to be less satisfying [65-68]. In other words, although applying standardized units, *the assays* are not standardized, and a test result of e.g. 1 µg/mL may be considered "therapeutic" in one assay, while "sub-therapeutic" in another assay. Official recommendations for drug levels one can consider "therapeutic" or "non-therapeutic" do not exist, leaving the decision to the individual clinician based on experience.

In ADAb assays, manufacturers usually provide a cut-off for "positive" and "negative", but the result will still have to be read on the background of the underlying method. Although drug-tolerant assays will detect ADAb with higher sensitivity, the clinical value of detecting low levels of ADAb is uncertain [61]. Differences in the ADAb-detecting sensitivity of different assays have led to wide ranges for the presence of ADAb to different TNF inhibitors. For example, for infliximab ADAb were present in 5% of patients in one study compared to 61% in

another [69, 70]. Most recently, panels of ADAb against infliximab and adalimumab with different binding characteristics have been developed [71]. This approach may contribute to a higher degree of assay standardization in the future.

In a clinical setting, direct comparisons of test results from different assays and interchangeable use of different modalities on the same patient/drug/ treatment course should currently be avoided.

It has been discussed how far drug levels alone are sufficient markers for patient-tailored treatment, because – independently of the presence of ADAb – drug levels within a therapeutic range can be considered sufficient, while drug levels below therapeutic range can be considered insufficient. Others have argued, however, that ADAb are necessary to decide on further treatment: dose escalation in case of low drug titers and negative ADAb versus switching to a different drug in case of low drug titers combined with ADAb. The presence of ADAb at an early stage of treatment may also be predictive for treatment responses at later stages [72].

Clinicians who want to be guided in their treatment decisions by drug levels and ADAb, have to become “specialists” on the applied assay, define which drug levels to consider as “therapeutic” and consider drug-tolerance and assay functionality and make sure that they test their patients at the right moment in the treatment interval.

In summary, drug level and ADAb measurements have been shown to correlate with treatment responses, they are cost-effective and clinically relevant [73-75], but leave room for misinterpretation due to different and not-standardized methodologies and recommendations, resulting in misguidance rather than guidance. Drug level and ADAb measurements may allow for a certain prediction of treatment responses within an ongoing treatment course [72]. But they cannot be used as tools of prediction before treatment with TNF inhibiting or other treatment strategies.

# 4. Identification of candidate biomarkers for TNF inhibitor responses by mass cytometry

## 4.1. Patient immune cells as reporters?

The reporter-gene assay discussed in the previous chapter reports TNF "activity" with the help of a genetically modified cell line. But can a patient's own cells be reporters of TNF activity?

The role of immune cell profiling in treatment stratification in rheumatic diseases was previously discussed by Ermann et al. [76], who also reviewed different methodological approaches for that purpose. Based on the role of TNF in RA, one expects higher TNF signaling activity in a majority of patients with active inflammatory disease, while patients successfully treated with a TNF inhibitor should express lower TNF signaling activity. Likewise, untreated RA patients with strong TNF signaling activity should be more likely to respond to treatment with a TNF inhibitor than those with initially low TNF signaling activity. Previous studies in flow and mass cytometry have indicated that RA patients indeed may express TNF signaling signatures, which differ from those of healthy individuals and patients with osteoarthritis [77, 78]. However, which signaling markers in which immune cell subsets are the best reporters for TNF signaling activity, and which methods are promising in the search for such candidate biomarkers? Can in-vitro stimulation of cells from patients and healthy donors with TNF contribute to an accentuation of pre-existing TNF profiles?

## 4.2. Introduction to mass cytometry

Mass cytometry was introduced in 2005 and commercialized from 2009, with the primary goal of satisfying the growing demand for "width *and* depth": the need to analyze large heterogeneous cell populations with a simultaneous read-out of multiple

parameters on a single cell level [79-81]. For the past decades, flow cytometry was the platform that filled the gap between these needs. The basic principles are the same for flow and mass cytometry: cellular epitopes are marked, usually with tagged antibodies as probes, and read out sequentially on a single cell basis. In flow cytometry, antibodies are tagged with fluorophores and read out in an optical system based on lasers and light spectra. Thus, flow cytometry has certain limitations concerning the total number of simultaneous markers, and complex compensation strategies have to be applied to handle spectral overlap. The problems of light spectrum limitations and spectral overlap are partially overcome by newer technologies such as “brilliant” dyes and spectral analyzers, but panels still rarely exceed 15-18 markers in flow cytometry, although panels of 30-40 markers technically are within reach [82]. Major advantages of flow cytometry are longstanding experience with almost unlimited access to validated antibodies, the collection of light scattering properties of a cell, the high throughput of cells, the recovery of (live) cells, allowing for cell sorting and subsequent analysis and, last but not least, the relatively low total cost.

Mass cytometry is a recent and developing technology. In mass cytometry, single cells are profiled applying metal-tagged antibodies against extra- and intracellular epitopes with a read-out in a mass spectroscopy time-of-flight chamber **(Figure 4)**. A limited, but growing range of ready-made metal-tagged antibodies are currently commercially available, other antibodies have to be metal-conjugated and validated by the user. The method contains no optical system and scatter properties of cells are therefore not collected. In preparation for read-out in the time-of-flight chamber, cells are ionized and therefore lost for further analysis or cell sorting. The cell throughput is lower than in flow cytometry and the total application cost higher. However, there are some major advantages to mass cytometry, mostly due to the use of rare earth metals (lanthanides) as probes and a mass spectroscopy time-of-flight chamber as detector. Metal isotopes used in mass cytometry have low natural abundancies and background “noise” is therefore much lower in mass compared to flow cytometry. Channel spillover is much lower than spectral overlap in flow cytometry and can be handled with thorough panel design and titration (see chapter 4.3.2.).

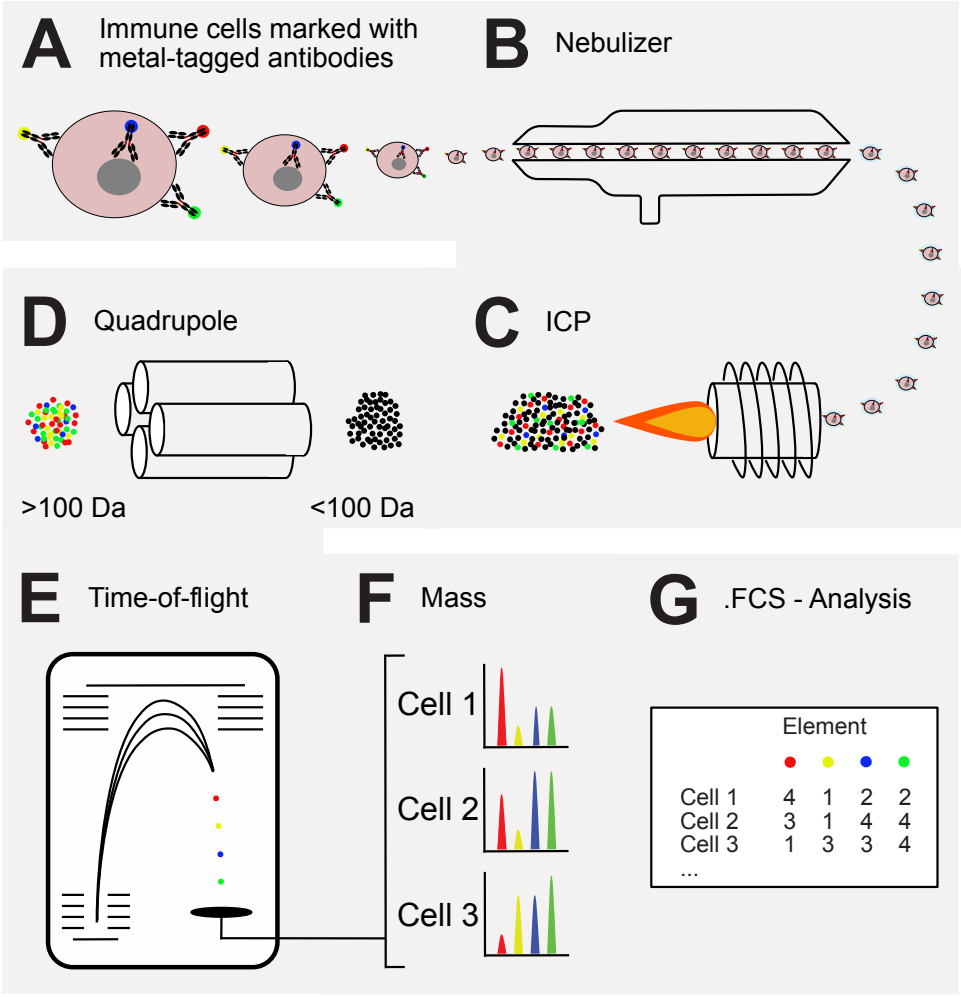



**Figure 4: Workflow mass cytometry.** Figure adapted from Bendall et al. [83]. Cells are marked with metal-tagged antibodies for extra- and intracellular epitopes **(A)**. Droplets including single cells are formed in a nebulizer **(B)** and injected into argon-plasma, where single cells are atomized and ionized **(C)**. Ions from the resulting ion-cloud are separated by size into smaller "debris" and reporter ions **(D)**. The latter are measured in a mass spectrometry time-of-flight chamber **(E-F)**. Collected data are stored as .FCS files **(G)** and ready for subsequent analysis.

In a mass cytometry experiment, usually some mass channels will be occupied for quality control such as standardization and normalization, e.g. the use of normalization beads [84]. Additionally, simultaneous measurements on samples from different donors, different conditions or the introduction of a reference sample require mass channels for multiplexing/ barcoding, where sample identities are specified

through unique combinations of e.g. six palladium isotopes $^{102}Pd$, $^{104}Pd$, $^{105}Pd$, $^{106}Pd$, $^{108}Pd$, $^{110}Pd$ or addition of a CD45 marker [85, 86].

## 4.3. Analysis of high-dimensional data

Mass cytometry data is high-dimensional. In each sample, up to millions of events are acquired, and each event carries individual data on multiple markers. Data analysis by traditional biaxial plotting in a hierarchical manner is not feasible for data sets of >20 markers per cell. For 20 markers, this would result in 190 plots to be analyzed, 30 markers would result in 435 plots for analysis. Two-dimensional gating is more likely to result in known cell populations, allowing comparisons with traditional flow cytometry data, but also more likely to oversee rarer or unknown cell populations defined by more atypical and unknown combinations of phenotyping markers.

A normal workflow in the analysis of mass cytometry data on immune cells will include:

1. a data clean-up with removal of doublets and debris
2. the definition of cell subsets based on phenotyping markers
3. analysis of functional markers within those cell subsets, in combination with clinical data

Several algorithms have been developed to cluster, embed and visualize multi-dimensional cytometry data as an alternative to traditional gating. In the case of a dataset with several individuals, data can be clustered/embedded simultaneously for all individuals. This approach is timesaving and provides a good general overview over e.g. a population, but not all algorithms can handle sets with millions of events from bigger cohorts. Also, differences in the co-expression of markers used for clustering/embedding, may lead to blurriness and a lower resolution at least for certain cell subsets. A more time-consuming approach would be to do clustering/embedding on datasets from each individual, providing a high resolution even for small cell subsets.

### 4.3.1. SPADE

*SPADE* (*S*panning-tree *P*rogression *A*nalysis of *D*ensity-normalized *E*vents) is a clustering algorithm, which seeks to provide a 2-dimensional visualization of high-dimensional data, usually based on phenotyping markers [87]. Single-cell resolution is here lost due to clustering. The SPADE algorithm performs a density-dependent down-sampling to avoid an underrepresentation of rarer cell types. This is followed by clustering of data of similar phenotypes and visualization in a spanning tree construction representing the original dimensionality of the data. The spanning tree of clusters can be overlaid with colors, e.g. indicating the fold change of a functional marker in each cluster **(**exemplary data shown in **Figure 5)**.

### 4.3.2. tSNE/viSNE

The primary purpose of *tSNE* (*t*-distributed *S*tochastic *N*eighbor *E*mbedding) is – similar to SPADE – to provide a visualization of high-dimensional data in two dimensions, but – in opposite to SPADE – with a preserved single cell resolution [88]. In cytometry, often an adaptation to tSNE, viSNE is used (*vi*sualization of t-distributed *S*tochastic *N*eighbor *E*mbedding). After an optional step of down-sampling, high-dimensional data is sorted by t-distributed stochastic neighbor embedding and spread out in two dimensions (tSNE1, tSNE2), retaining as much as possible of the high-dimensional information for each event. Resulting biaxial scatter plots are called viSNE maps. viSNE maps are typically based on the expression of phenotyping markers. In a further step, color overlays representing the expression of functional markers can be made **(**exemplary data shown in **Figure 5)**.

### 4.3.3. CITRUS

The *CITRUS* algorithm (Cluster identification, characterization, and regression) combines clustering and regression modeling to identify correlative or predictive markers for differences between two or more cohorts of individuals [89]. Although not primarily designed for the evaluation of paired samples – e.g. from the same patients before and after a certain treatment – the algorithm is applied for this purpose as well.

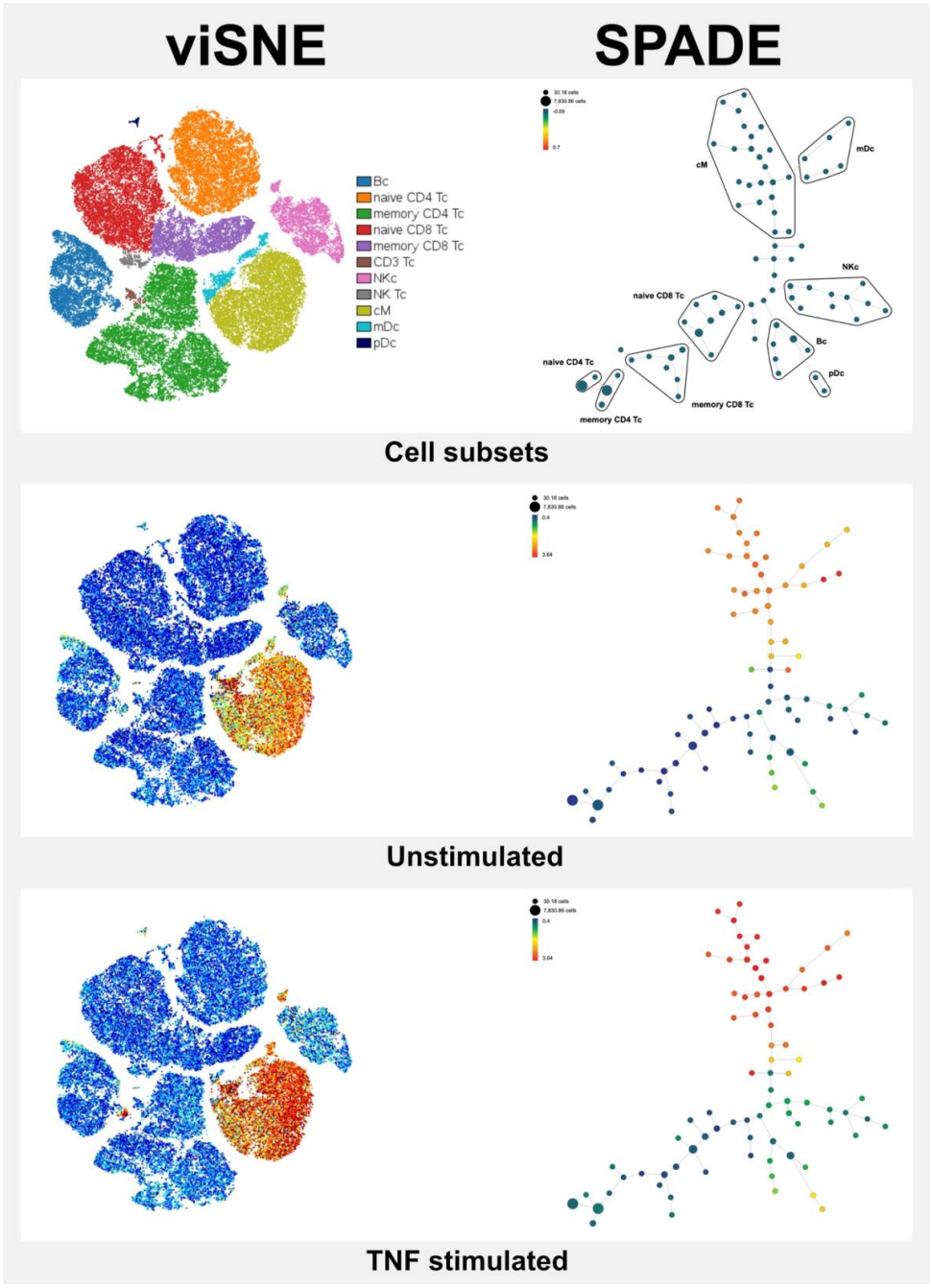



**Figure 5: Examples of viSNE (left) and SPADE (right)** on 50.000 PBMC from one RA patient (own data), gated into cell subsets (upper). The same plots and population distributions with color-overlay for the expression of p-p38 in unstimulated (mid) and TNF stimulated PBMC (lower).

CITRUS does not include a data clean-up, which has to be performed manually before feeding data for analysis. The algorithm contains an optional down-sampling step, in which cell numbers can be adjusted so that individuals are represented with equal amounts of cells. All cells are then clustered after phenotypes, and clusters can be analyzed either by differences in abundance based on phenotyping markers or by differences in the median expression of functional markers. Both correlative and predictive models can be applied, the latter with cross validation.

### *4.3.4. NM2B*

*Individualized analysis workflow "NM2B".* For the analysis of data for paper 3 we developed an individualized algorithm to fit better the underlying data and research question. This algorithm covers all steps from data clean-up, through finding cell types to classification of patients versus healthy donors utilizing data from both unstimulated and TNF-stimulated conditions.

The approach for data clean-up is based on a talk and follow-up publications by Bruce Bagwell, where he describes the use of the parameters "Event_length", "Center", "Offset", "Width", "Residual", "191Ir_DNA1", "193Ir_DNA2" as markers for identifying live single cells [90, 91]. Bagwell gates manually on a tSNE plot for identifying cell populations. Since clustering on a tSNE plot is in general not recommended, a different approach is applied in this algorithm, where cells are clustered based on these markers directly in the 7-dimensional space. Mean-variance scaling for each of the markers is used in order to get comparable features. A Gaussian mixture model of two multidimensional Gaussian distributions by expectation maximization is then applied and used for clustering.

Our approach for finding cell types is based on splitting the dataset up into many small clusters and then combining such small homogeneous clusters into meta-clusters. All data is arcsinh-transformed with a cofactor of 5 as suggested by Qui [92]. Farthest-point sampling is a fast algorithm to cluster cells by beginning with an initial point and then selecting the point that is farthest away from the closest of all previously selected points until one has the desired number of clusters. These selected

points are then considered as the cluster centers and all remaining points are assigned to the closest cluster center. This procedure can be shown to approximate k-means clustering. When splitting the data into enough clusters in this way, this results in relatively small homogeneous clusters.

For classification, a set of classification markers (“functional markers”) is defined. For each meta-cluster the median and 90% quantile of each of the functional markers for all basal cells is calculated. In addition, arcsinh ratios of the expression of functional markers in stimulated and basal cells are calculated. Three models are tested, based on either only basal variables (basal), or only arcsinh ratios between stimulated and basal variables (ratio) or both basal and arcsinh ratios (combined). A logistic lasso regression model with automatic variable selection is fitted, using double leave-one-out cross validation. The algorithm reports cross-validation accuracy, area under the ROC curve (AUC), and all non-zero coefficients.

# Objectives

The clinical challenge we are facing is a lack of reliable biomarkers for patient classification and stratification in RA.

Objectives of the work underlying this thesis were:

1. **To explore and compare existing markers for TNF inhibitor drug monitoring**. For that purpose, we compared three existing assays for the measurement of drug trough levels and anti-drug antibodies on serum samples from 107 patients (**Paper 1**).
2. **To collect patient material and data, and to set up a methodological background for mass cytometry experiments**. For that purpose, PBMC from RA patients and healthy donors were cryopreserved, and clinical data were collected. A smaller set of quality control experiments were performed, addressing signaling responses in PBMC, depending on bench-time before cryopreservation, culturing time after cryopreservation and TNF cytokine titrations in order to establish adequate stimulation time and dose for TNF (all unpublished data). An antibody panel for mass cytometry experiments on cells from RA patients and healthy donors was designed and titrated. A hierarchical panel titration methodology based on a backbone panels and individual subpanels was established and published (**Paper 2**).
3. **To explore immune cell subsets for signaling signatures in RA patients compared to healthy individuals, with a primary focus on TNF signaling**. For that purpose, we compared signaling signatures in unstimulated and TNF stimulated PBMC from 20 newly diagnosed RA patients and from 20 healthy donors (**Paper 3**).

# Results

## 1. Exploration and comparison of existing markers for TNF inhibitor drug monitoring (**paper 1**)

Immunogenicity is a frequent cause of secondary non-response to TNF inhibitors. Drug level measurement and detection of ADAb have been shown to be cost-effective and clinically relevant, and a large number of assays are available for

| | |
|---|---|
| Number of patients | 107 |
| Female/male | 53/54 |
| Mean patient age (min-max) | 51 years (16-86) |
| Diagnoses, number of patients (% of total) | |
| Rheumatoid arthritis | 37 (34.6) |
| Spondyloarthritis (excl. psoriatic arthritis) | 18 (16.8) |
| Psoriatic arthritis | 17 (15.9) |
| Juvenile idiopathic arthritis | 9 (8.4) |
| Psoriasis | 23 (21.5) |
| Others | 3 (2.8) |
| Mean disease duration (min-max) | 19 years (2-50) |
| Medication, infliximab (IFX) | |
| IFX-Remicade (%) | 94 (88) |
| IFX-Remsima (%) | 13 (12) |
| Median treatment duration (min-max) | 70 months (0-158) |
| Mean IFX dose (min-max) | 4.5 mg/kg (2.1-10.4) |
| Mean IFX dose rheumatology (min-max) | 3.8 mg/kg (2.1-7.3) |
| Mean IFX dose dermatology (min-max) | 6.4 mg/kg (4.5-10.4) |
| Mean IFX interval (min-max) | 7.5 weeks (4-17) |
| Co-medication | |
| Methotrexate | 76 patients |
| Other DMARD | 2 sulfasalazine, 1 leflunomide, 1 azathioprine |
| Prednisolone | 8 patients |

**Table 2: Patients.** Patient characteristics, diagnoses, medication and co-medication.

these purposes. It is, however, difficult to compare assays and translate results into clinical meaningful information due to different methodological approaches and a lack of assay standardization.

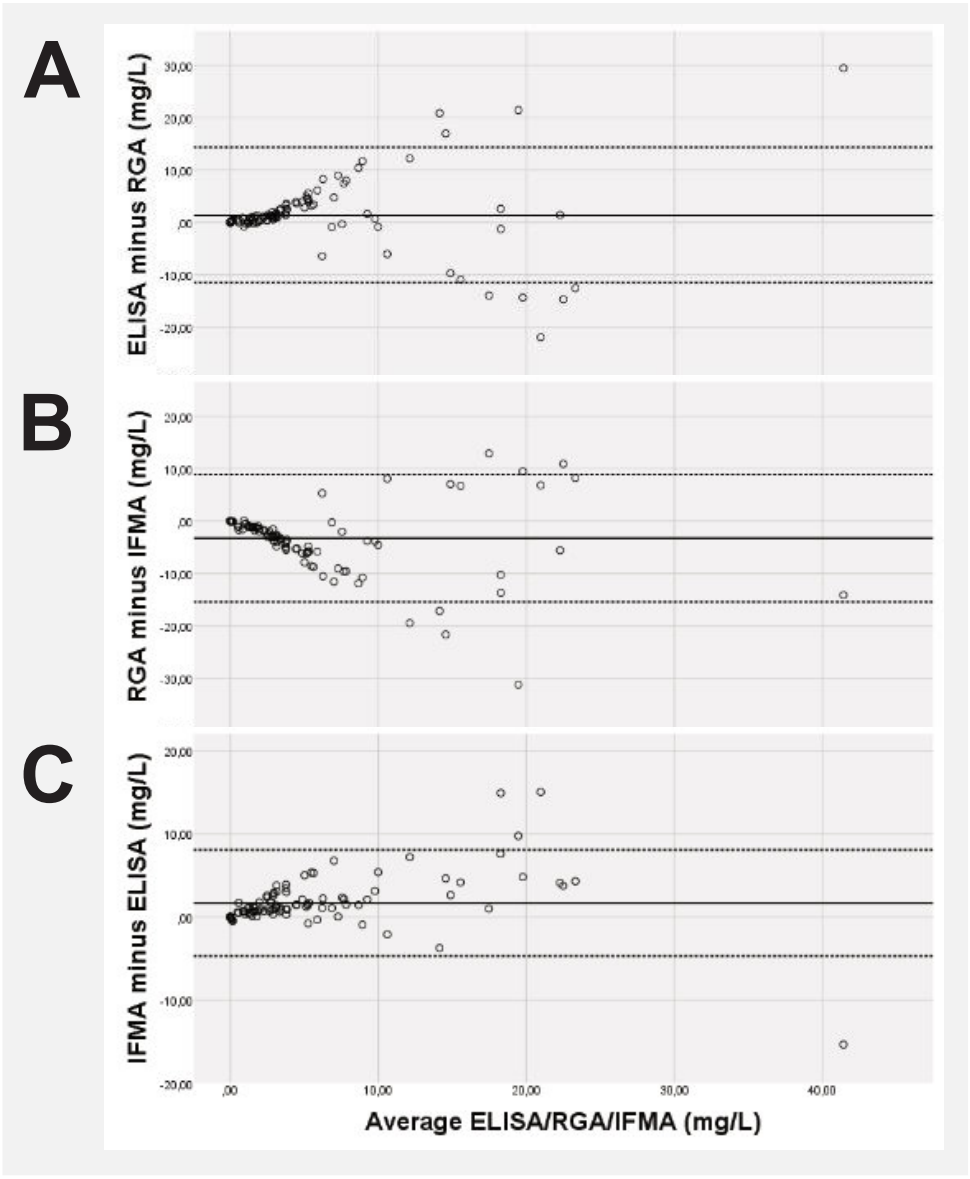



**Figure 6: Agreement between assays for drug levels.** Bland–Altman plots demonstrating agreement between IFX levels determined using different techniques. The average result from all three assays for one patient sample (x-axis) is plotted against the difference between the two assays compared (y-axis) for

**A**: ELISA versus RGA,

**B**: RGA versus IFMA, and

**C**: IFMA versus ELISA.

The difference is 0 in case of total agreement; increasing distance from 0 indicates an increasing disagreement between the two compared assays.

To explore and compare existing markers of TNF inhibitor drug monitoring, we analyzed infliximab drug levels and antidrug antibodies in 107 patient samples **(table 2)** using ELISA and RGA, as previously described **(Figure 3)**. Additionally, immunofluorometric assays (IFMA, in-house method at Oslo University Hospital) were carried out as described in [93].

The three assays were in better agreement at lower IFX levels and in poorer agreement for higher levels. No samples with an average IFX level <10 mg/L were outside the limits of agreement, which were narrowest for ELISA versus IFMA. The

RGA resulted in systematically lower IFX levels than the ELISA, whereas the IFMA resulted in higher levels than the ELISA **(Figure 6A-C)**.

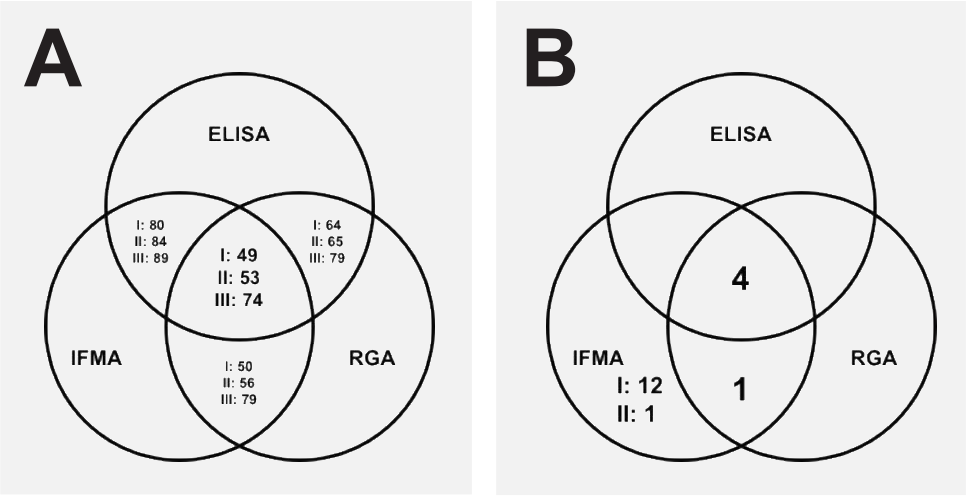



**Figure 7: Agreement between assays for ADAb.** Venn diagrams illustrating agreement for categorical results in ELISA, IFMA and RGA.

**A**: Percentages of samples which end up in the same categories for IFX levels in all assays (center) and in two assays (sides). I: therapeutic range of 3-8 mg/L for all assays; II: therapeutic range of 1.5-12 mg/L for all assays; and III: individualized ranges of RGA 1-9 mg/L, ELISA 2-10 mg/L and IFMA 3-11 mg/L

**B**: The number of ADA-positive patient samples (of the total of 107) are shown as follows: four samples were detected as positive all assays, 1 sample was positive in the RGA and the IFMA, 12 samples were positive only by IFMA. Removal of all 'low-positive' samples from IFMA resulted in one positive sample only detected by the IFMA and an overall agreement on the outcome of 98% of samples.

We evaluated several therapeutic drug ranges **(Figure 7A)**. Application of two fixed therapeutic ranges for all three assays (one narrower, one wider) resulted in low rates of agreement for about 50% of the patient samples, whereas an individual adaptation of categories to each assay improved the rate of agreement to 74% of all patient samples. For all categories, agreement rates were significantly higher between ELISA and IFMA than ELISA and RGA or IFMA and RGA.

Of the 107 patient samples, four were IFX ADAb positive in the ELISA (≥5 AU/mL). The same four samples were found to be positive for ADAb by the RGA, as was one additional sample that was just below the threshold in the ELISA (4.7 AU/mL). The IFMA identified a total of 17 ADAb-positive samples, 11 of which fell into the category 'low positive' (10-30 AU/L). The remaining six samples included the ADAb-positive samples from ELISA and RGA. One patient, who was classified as IFX ADAb-positive in all three assays, was treated with biosimilar IFX Remsima®. Agreement for categorical results for ADAb from all three assays is presented in **Figure 7B**.

In conclusion, we found that TNF inhibitor monitoring assays measure on different scales and that the agreement between quantitative results is limited. However, inter-assay differences could partially be overcome by assay-individualized translations of quantities into categories, which is also necessary for a meaningful clinical application. Our data demonstrate that assays should not be used interchangeably, and that direct comparisons of quantitative drug levels obtained with different assays should be avoided.

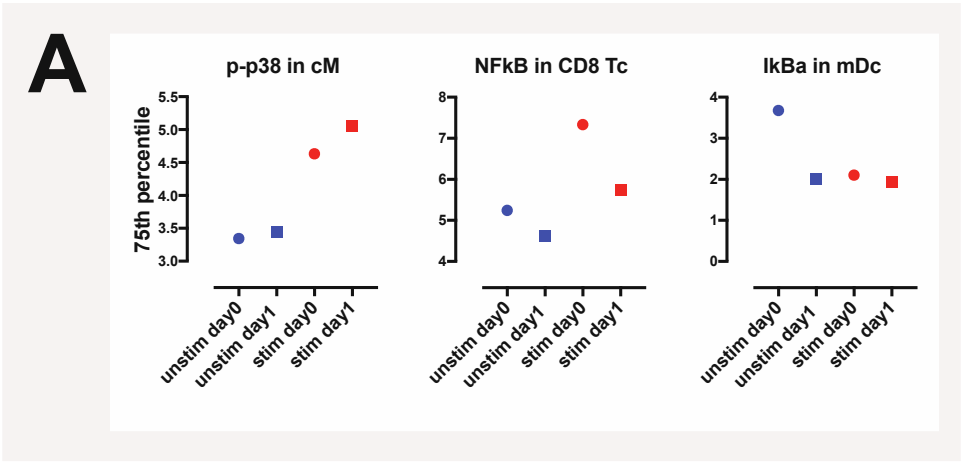
# 2. Collection of patient material and data, and set-up of a methodological background for mass cytometry experiments (unpublished and **paper 2**)

## 2.1. Collection and storage of cells

Two common methods for the collection and storage of immune cells from peripheral blood have been used in the work underlying this thesis. After venipuncture, blood can undergo density gradient centrifugation, which separates red blood cells, platelets and granulocytes from the reminder of white blood cells (predominantly lymphocyte, monocyte, dendritic cell and natural killer cell lineages). The latter, the so-called peripheral blood mononuclear cells (PBMC), are then cryopreserved "alive" either in liquid nitrogen or in an ultra-deep freezer at around -150 °C. Alternatively, whole blood undergoes immediate fixation. Before or after storage at -80 °C, erythrocytes are removed by lysis, resulting in the so-called peripheral blood cells (PBL). The same techniques as described for PBMC and PBL can be applied to bone marrow (BM).

## 2.2. Quality control and TNF titration

In smaller sets of quality control experiments, we addressed questions regarding signaling responses in PBMC. First, we explored the impact of bench-time before cryopreservation on signaling in PBMC. Blood was drawn from three healthy donors, and half of the material was cryopreserved immediately, while the other half was kept in cell preparation tubes (CPT™) after density centrifugation at room temperature and cryopreserved on the next day (bench-times up to 24 hours are acceptable following the manufacturer). After thawing, all samples were analyzed simultaneously regarding the expression of functional markers with and without TNF stimulation. We found that basal signaling for some markers as well as responses to TNF stimulation were weaker, while cisplatin, a marker for cell death, was increased



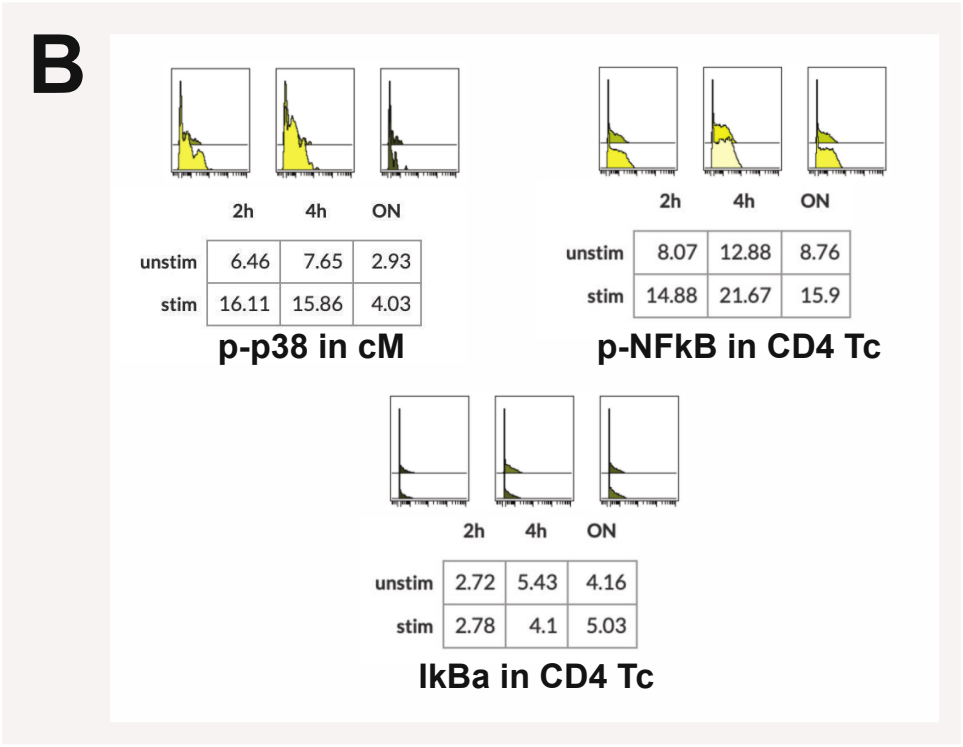




**Figure 8: QC experiments.**

**A**: Exemplary data from bench-time experiments. Expression of functional markers in specific cell subsets (means, N=3) for PBMCs processed on the day of sample acquisition (day0) or after one day of bench-time (day1).

**B**: Exemplary data from resting time experiments, showing the expression of functional markers in specific cell subsets (N=1), when rested after thawing for 2 hours (2h), 4 hours (4h) or overnight (ON).

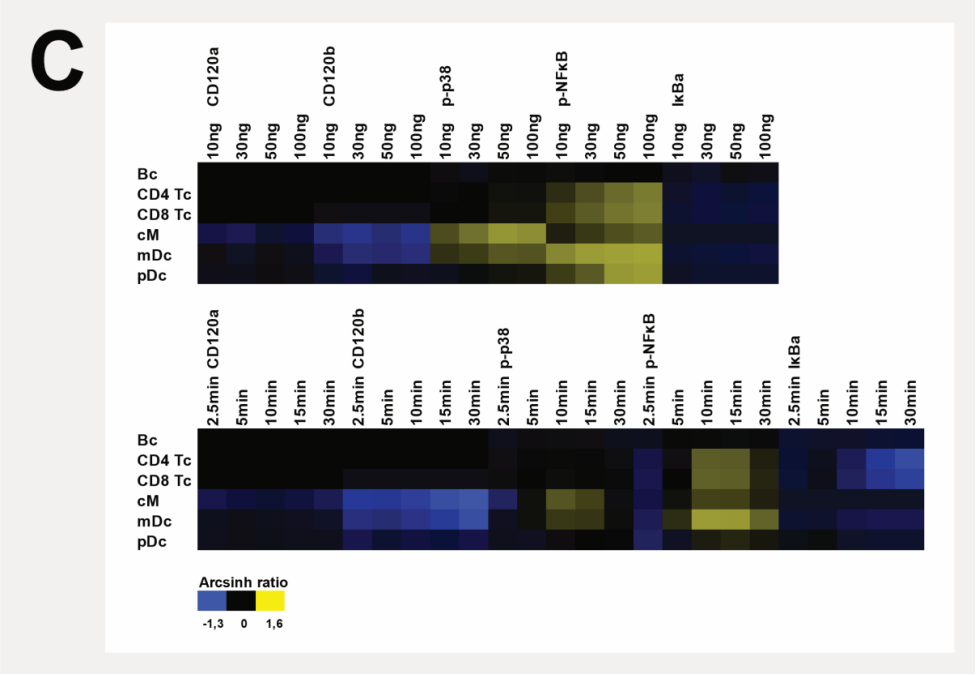



**Figure 8: QC experiments.**

**C**: TNF titration experiments on exemplary cell subsets and functional markers for TNF dose with a fixed stimulation time of 15 minutes **(upper)** and TNF time with a fixed stimulation dose of 50 ng **(lower)**.

in PBMC with long bench-time compared to those cryopreserved immediately. We concluded that only PBMC which had been cryopreserved on the same day of sample acquisition could be used for signaling studies (exemplary data shown in **Figure 8A**). Second, we explored the effects of different resting times after cryopreservation on signaling in PBMC. For that purpose, PBMC from two healthy donors were thawed, followed by resting for either two hours, four hours or overnight. All samples were analyzed simultaneously regarding the expression of functional markers without and with TNF stimulation. We concluded that a resting time of four hours provided the best conditions regarding signaling responses and cell death **(exemplary data shown in Figure 8B)**. Third, TNF dose and time titrations were performed to establish optimal stimulation conditions, resulting in TNF doses of 50 ng/mL and TNF stimulation times of 12-13 minutes for optimal signaling responses **(exemplary data shown in Figure 8C)**.

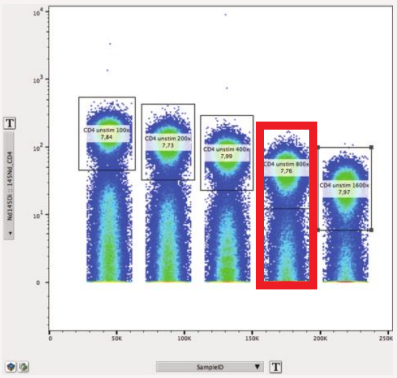
## 2.3. Panel design and antibody titration

Antibody panels for mass cytometry face three basic challenges: (1) non-specific antibody binding, which is an issue for any antibody-based assay, (2) signal spillover, which can be especially problematic in combination with (3) variations in sensitivity throughout the mass spectrum of the cytometer.

Although signal spillover from one mass channel to one or several other mass channels does not reach the same dimensions as spectral overlap in flow cytometry and although "compensation" as known from flow cytometry is unnecessary in mass cytometry, there are good reasons for thorough panel design and the use of pre-titrated antibodies. Mass cytometry has a lower sensitivity compared to flow cytometry, which is of concern especially for markers with low abundancy such as e.g. intracellular phospho-epitopes [94]. For these markers, even smaller amounts of signal spillover can be relevant. Besides non-specific antibody staining as an issue related to all antibody-based immunoassays, there are several known sources of signal spillover in mass cytometry. First, there can be spillover of isotopes from one mass channel into the adjacent +1 mass channel, e.g. from Sm149 to Nd150. Signal spillover is also stronger between different isotopes of the same metal. As an example, seven isotopes of the lanthanide neodymium are used in mass cytometry (Nd142, 143, 144, 145, 146, 148, 150), and spillover is more likely between these. A third source of spillover arises through oxidation of metal isotopes. One oxygen-atom has a molecular mass of 16; oxidation of a metal isotope will therefore result in spillover in the +16 channel. Oxidized Nd142 will followingly be detected as Gd158.

In panel design, markers of expected lower abundancy will therefore preferably be placed (1) in the middle of the mass range due to highest sensitivity of the detector in that range and (2) in channels with expected low spillover, especially from highly abundant markers (+1, +16 or isotope spillover). A free online tool (http://www.dvssciences.com/mydvs/) and instruction manual (http://www.dvssciences.com/mydvs/Maxpar_Panel-Designer_ug_100-9557A2_150925.pdf) for panel design is offered by Fluidigm for registered users.

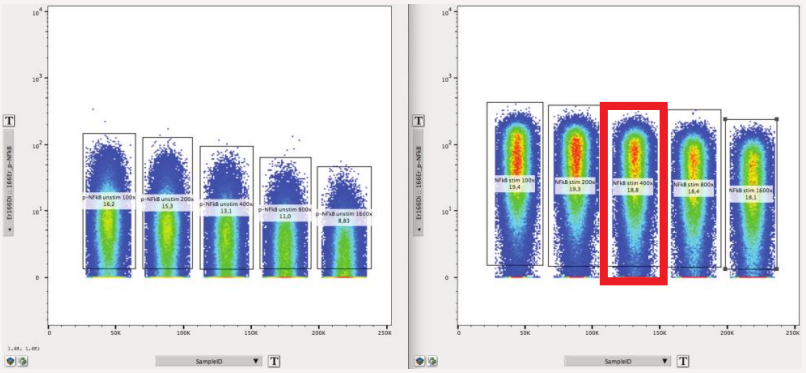
**A**



75th percentile

| Purpose | Target | Metal | 100x | 200x | 400x | Fluidigm | 800x | 1600x |
|---|---|---|---|---|---|---|---|---|
| Nd | Caspase3 | 142Nd | 0 | 0 | 0 | | 0 | 0 |
| | CD4 | 145Nd | 193 | 138 | 93,3 | 70 | 57 | 31,5 |
| +1, Nd | CD8a | 146Nd | 2,91 | 1,55 | 1,34 | | 0 | 0 |
| Nd | CD16 | 148Nd | 0 | 0 | 0 | | 0 | 0 |
| | Tolerance | | 38,6 | 27,6 | 18,66 | 14 | 11,4 | 6,3 |

**B**



75th percentile

| Purpose | Target | Metal | 100x | 200x | 400x | Fluidigm | 800x | 1600x |
|---|---|---|---|---|---|---|---|---|
| -1 | CD120a | 165Ho | 0 | 0 | 0 | | 0 | 0 |
| | p-NFκB uns | 166Er | 11,6 | 9,33 | 7,46 | 7 | 6,16 | 5,14 |
| | p-NFκB stim | 166Er | 84,4 | 83,7 | 70,3 | | 55,9 | 43,5 |
| | Ratio | | 7,275862 | 8,971061 | 9,423592 | | 9,074675 | 8,463035 |
| +1/Er | p-cJun | 167Er | 0,2 | 0,18 | 0 | | 0 | 0 |
| | Tolerance | | 2,32 | 1,866 | 1,492 | 1 | 1,232 | 1,028 |

**Figure 9: Antibody titration** exemplified on the phenotyping marker CD4 145Nd **(A)** and the functional marker p-NFkB 166Er **(B)**. Concatenated files and tables listing 75th percentiles of dual count expression for each antibody concentration (100x, 200x, 400x, 800x, 1600x). The tables list covers also channels of potential spillover (+16 not shown due to irrelevance for the actual panel). Red boxes indicate chosen antibody concentrations. As a reference, marker expression and tolerance as measured by the vendor Fluidigm are listed (green background).

Primary antibody panels were designed for QC experiments and the main experiment (**Paper 3, Figure 9**). These panels were partially based on immune phenotyping panels offered by the vendor Fluidigm. In parallel, we worked on a hierarchical panel titration method for more complex mass cytometry panels (**Paper 2**), as presented in the following.

## 2.4. Hierarchical approach to panel titration

We developed a hierarchical approach for mass cytometry panel titration – exemplified on peripheral blood (PB) and bone marrow (BM) cells – which is suitable for complex mass cytometry panels and takes into account both spillover and non-specific binding of antibodies. In each titration step, channels which potentially may receive spillover (±1 Da, +16 Da and channels detecting isotope impurities) were kept empty. The optimal titer for each antibody was approximated by balancing the ability to discern positive from negative cells with the amount of spillover generated in other mass channels. In titration step 1, a simpler *backbone* panel was titrated and optimized to provide a wide hematopoietic background **(Figure 10A)**. On the background of this backbone panel, subpanels consisting of further surface antibodies for deeper characterization (for example CD45RO expression on CD4 Tc) were added in titration step 2 **(Figure 10B)**. A singular SPADE clustering was performed on the entire data set including all titrations (6.4x10^6 single cells) to manually identify cell subsets based on the backbone panel **(Figure 10C)**. The signal from the titrated antibodies was then measured separately in each of the cell subsets, plotted in heatmaps and displayed side-by-side as a function of antibody dilution. Antibody titers were evaluated based on their expression on relevant cell subsets, e.g. CD45RO on T helper cells **(Figure 10D)**. In a third titration step, markers for signaling were added to cells after ex-vivo stimulation with GM-CSF, IFNα, LPS – included an unstimulated control – with a rich phenotypic background consisting of panels from both titration step 1 and 2 **(Figure 10E)**.

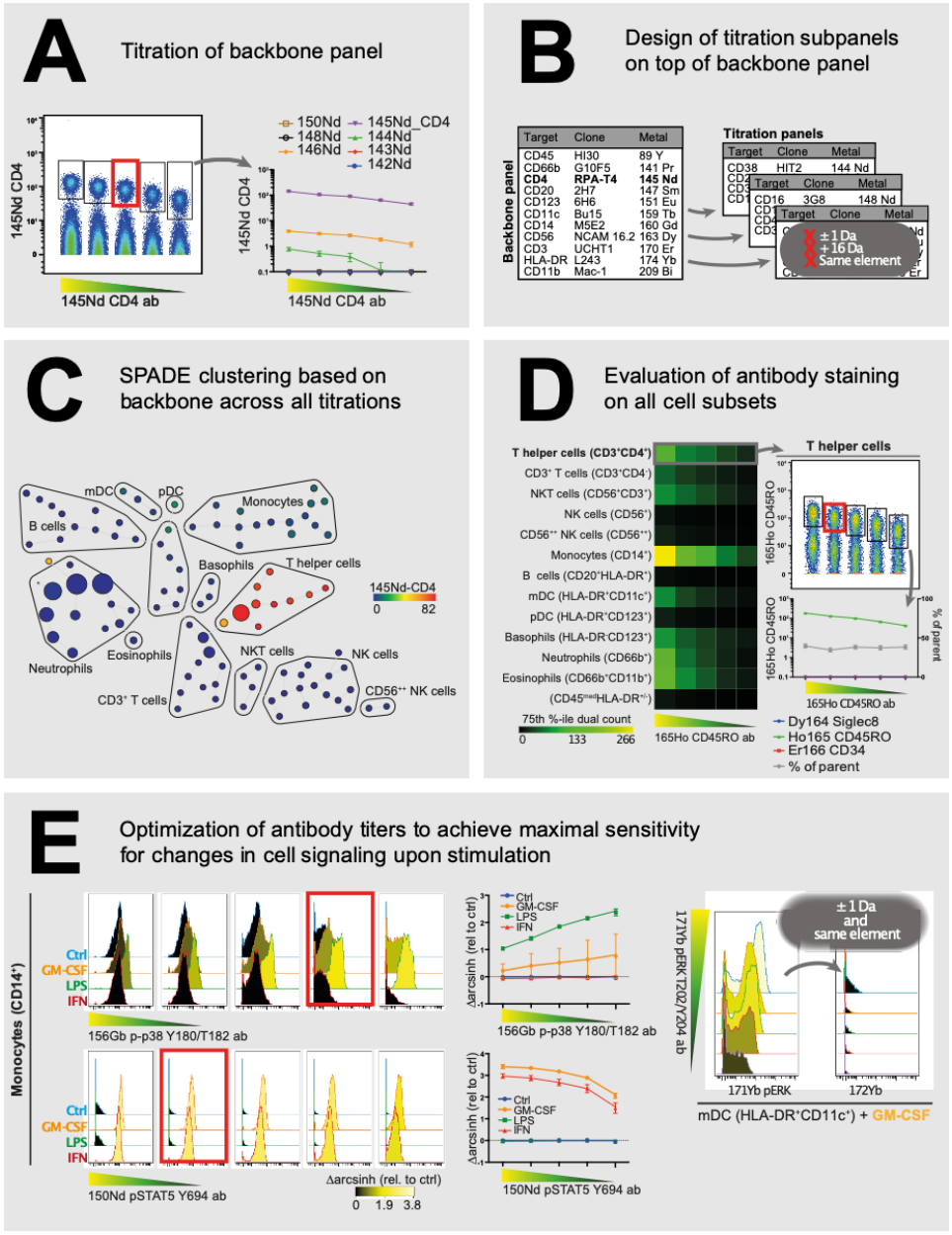



**Figure 10: Workflow for hierarchical antibody titration and panel design.**

**A:** Titration step 1: backbone panel providing a wide hematological background

**B:** Titration step 2: deeper characterization of cell subsets by addition of further surface markers

**C:** SPADE on cells across all titrations using the backbone panel for clustering

**D:** Exemplary data for the marker CD45RO, titrated on the relevant cell subset of T helper cells

**E:** Titration step 3: signaling panel titration after stimulation with GM-CSF, LPS and IFNα (left). Example of spillover from 171Yb pERK into 172Yb (right).

# 3. Exploration of immune cell subsets for signaling signatures in RA patients compared to healthy individuals, with a primary focus on TNF signaling (**paper 3**).

The pro-inflammatory cytokine tumor necrosis factor (TNF) plays a central role in the pathogenesis of rheumatoid arthritis (RA), and TNF inhibitors constitute an efficient treatment against inflammatory activity in RA. The objective of this study was to use unbiased methods to identify signaling patterns in immune cell populations from RA patients with an emphasis on TNF signaling. We employed mass cytometry (CyTOF) with a panel of 13 phenotyping and 10 functional markers to compare signaling signatures in unstimulated and TNF-stimulated peripheral blood mononuclear cells (PBMC) from 20 newly diagnosed, untreated RA patients and 20 healthy donors. The resulting high-dimensional data were analyzed in three independent analysis pipelines, characterized by differences in both data clean-up, identification of cell subsets/clustering and statistical approaches.

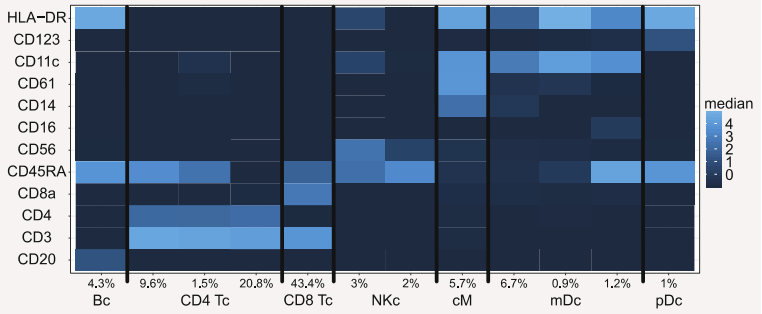
In analysis pipeline 1 (NM2B algorithm), single-cell data from all 40 individuals were clustered and meta-clustered, and different numbers of clusters and meta-clusters were tested. The model used provided the best translation of meta-clusters into common immune cell subsets. Results presented here are based on 49 clusters and 12 meta-clusters; the latter include one B cell meta-cluster (4.3%), four of T cells (75.3%), two of natural killer cells (5%), one of classical monocytes (5.7%), three of myeloid dendritic cells (8.8%) and one of plasmacytoid dendritic cells (1%). Phenotyping markers are differentially expressed in the meta-clusters **(Figure 11A)**; differences in expression of phenotyping markers in healthy donors vs. RA patients were not significant.

A regression model based on both basal expression of functional markers and arcsinh ratios (“combined model”) provided the best predictive TNF signaling patterns for healthy donors and RA patients. In this model seven functional markers (IkBa,

CD120b, CD86, p-cJun, p-NFkB, p-p38, and p-Akt) in five cell subsets (memory CD4 Tc, two mDc subsets (HLA-DR+CD14low and HLA-DRhigh CD14-), naïve CD4 Tc (CD11clow), and cM) were identified as predictive markers **(Figure 11B)**. Applying these markers, the combined model correctly classified 18 of 20 RA patients and 17 of 20 healthy donors **(Figure 11C)**. Principle component analysis (PCA) of features identified by the Lasso-regression showed a good separation of HD vs. RA in the combined **(Figure 11D)** and basal model, but to a lesser degree in the ratio model.

**A**

12 meta-clusters (vertical columns), divided into 7 common immune cell subsets. Heatmap over the expression of phenotyping markers in each meta-cluster




**B**

Results from Lasso regression: predictive features (functional markers and cell subsets) for the classification of HD and RA (combined model)

| Marker | Cluster (number) | Quantile | Variable | Coef |
|---|---|---|---|---|
| IkBa | memory CD4 Tc (4) | q90 | basal | -6.01 |
| CD120b | mDc (9) | q90 | ratio | -5.24 |
| p-cJun | naïve CD4 Tc (3) | median | ratio | 4.72 |
| CD86 | mDc (9) | median | ratio | -4.63 |
| CD120b | naïve CD4 Tc (3) | median | ratio | 3.54 |
| CD86 | cM (8) | median | basal | 2.43 |
| p-NFkB | mDc (9) | median | basal | -2.08 |
| p-p38 | memory CD4 Tc (4) | median | basal | -1.84 |
| p-Akt | naïve CD4 Tc (3) | q90 | ratio | 0.22 |
| IkBa | mDc (10) | q90 | basal | 0.12 |
| CD86 | mDc (10) | median | basal | 0.07 |

**Figure 11A-B: Results NM2B algorithm.**

**A:** Heatmap over the expression of phenotyping markers in 12 meta-clusters (columns). Meta-clusters were identified as B cells (Bc), CD4+ T cells (CD4 Tc), CD8+ T cells (CD8 Tc), natural killer cells (NKc), classical monocytes (cM), myeloid dendritic cells (mDc), and plasmacytoid dendritic cells (pDc). Relative abundance is given for each cell subset in percent.

**B:** Results from Lasso regression: predictive features (functional markers and cell subsets) and their contribution to the classification of healthy donors (HD) and RA patients (RA). Only nonzero coefficients are shown. Coefficients for CD86 and p-cJun are based on ratios and therefore inverted compared to CITRUS and manual comparisons of basal marker expression.

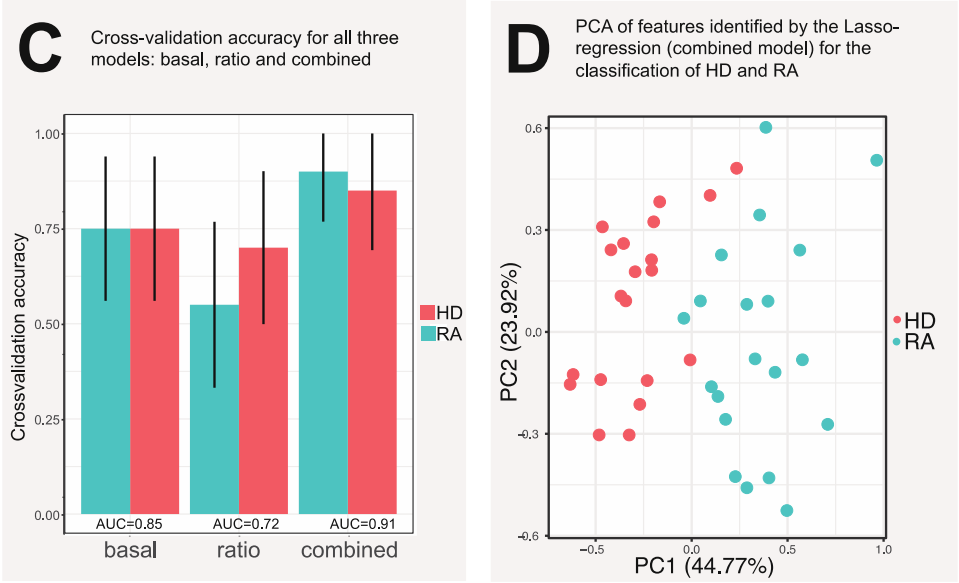



**Figure 11C-D: Results NM2B algorithm.**

**C:** Cross-validation accuracy for all three NM2B analyses ("basal," "ratio," and "combined"), with area-under-the-curve (AUC) values for ROC analysis.

**D:** Principle component analysis (PCA) of features identified by Lasso-regression (combined model) for the classification of RA patients and healthy donors.

In analysis pipeline 2 (CITRUS algorithm), we performed four repetitive CITRUS analyses of basal expression of functional markers. p-p38, IkBa, p-cJun, p- NFkB, and CD86 were identified as predictive markers by CITRUS, with memory CD4+ T cells being the most relevant cell subset for both p-p38, IkBa, and p-cJun, while clusters within myeloid dendritic cell subets (mDc) and classical monocytes (cM) were the most relevant for p-NFkB and CD86. There was not always a clear distinction between mDc and cM in hierarchical clustering in CITRUS, and both these cell subsets were relevant for the markers p-NFkB and CD86 **(Figure 11E)**. Compared to the NM2B and manual analysis, there were slight differences in the weighting of cell subsets for p-cJun. In four CITRUS analyses, memory CD4 Tc were the primary cell subset of interest for p-cJun, whereas automated analysis pointed to naïve CD4 Tc as the most significant cell subset. In manual analysis, p-cJun expression was significantly different between HD and RA in both naïve and memory CD4 Tc.

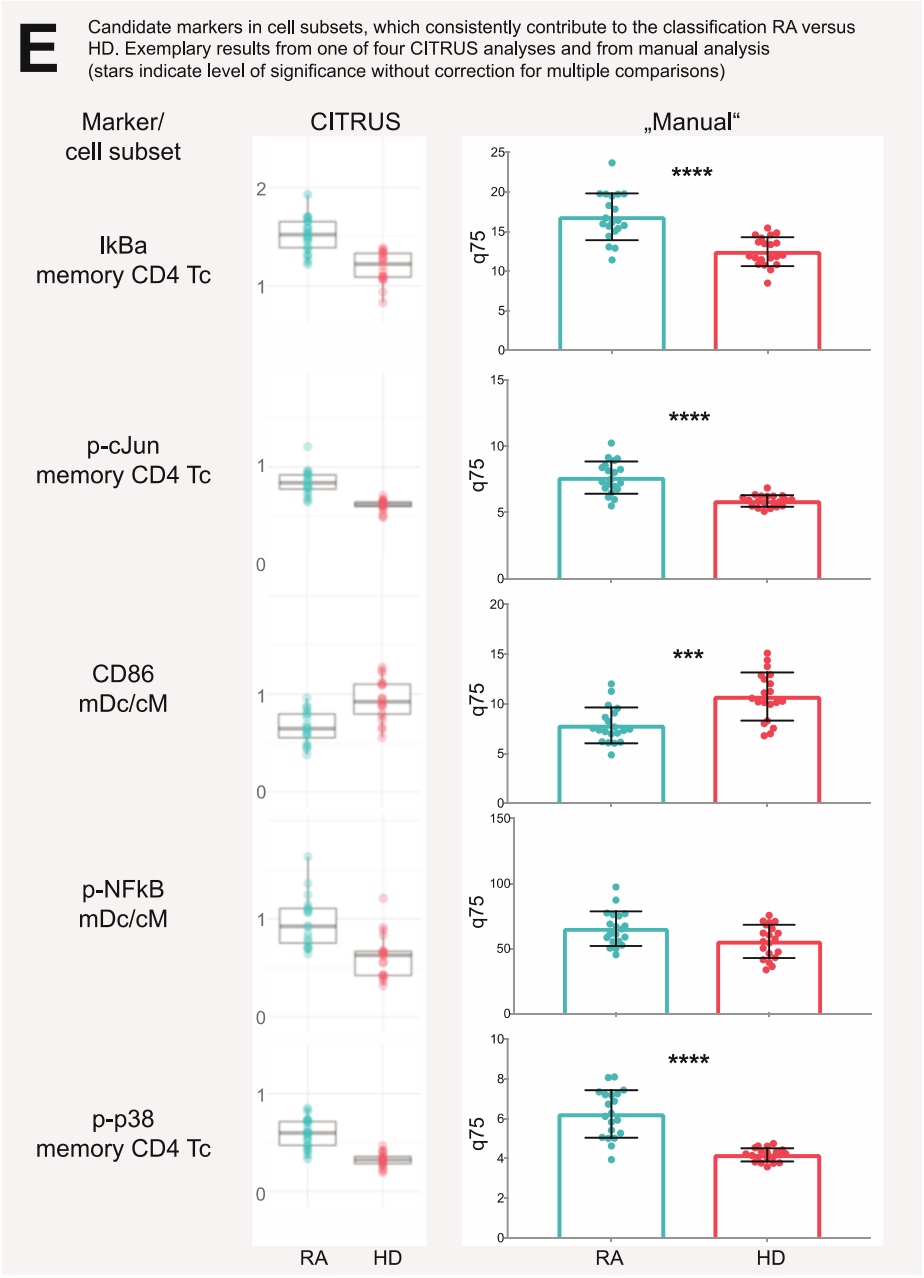



**Figure 11E: Results from CITRUS and manual analysis**

**E:** Cluster identification, characterization, and regression algorithm (CITRUS) and non-parametric testing (Manual). CITRUS results are presented in boxplots, as provided by the algorithm. Results from manual analysis are presented in scatter dot plots. Medians (CITRUS) and 75th percentiles (manual) are plotted for each RA patient (blue) and healthy donor (red); median, and upper and lower quartile. Asterisks indicate level of significance without correction for multiple comparisons (***$p < 0.001$, ****$p < 0.0001$).

Analysis pipeline 3 (manual analysis) generally confirmed results from regression data modeling **(Figure 11E)**. For the p-NFkB, regression tools had suggested significant differences in myeloid dendritic cell and classical monocyte subsets. However, in non-parametric testing, p-NFkB expression was not significantly different in myeloid dendritic cell and classical monocyte subsets of healthy donors vs. RA patients, while being significantly different in memory CD4 Tc.

In conclusion, all three analysis pipelines identified p-p38, IkBa, p-cJun, p-NFkB, and CD86 in cells of both the innate arm (myeloid dendritic cells and classical monocytes) and the adaptive arm (memory CD4 Tc) of the immune system as markers for differentiation between RA patients and healthy donors. Inclusion of the markers p-Akt and CD120b resulted in the correct classification of 18 of 20 RA patients and 17 of 20 healthy donors in regression modeling based on a combined model of basal and TNF-induced signal. Expression patterns in a set of functional markers and specific immune cell subsets were distinct in RA patients compared to healthy individuals. These signatures may support studies of disease pathogenesis, provide candidate markers for response, and non-response to TNF inhibitor treatment, and aid the identification of future therapeutic targets.

# Discussion

RA is a heterogeneous disease with a plethora of treatment options, and patients will profit from more exact classification and stratification. The work presented here highlights the current lack of classification and stratification markers and shows how single cell profiling by mass cytometry may contribute to the search for such markers. Methodological aspects such as antibody panel design and approaches for the analysis of high-dimensional data are emphasized.

## 1. Existing markers for TNF inhibitor drug monitoring

We first compared existing assays for TNF drug levels and ADAb and found that TNF inhibitor monitoring assays measure on different scales and that the agreement between quantitative results is limited. Inter-assay differences could partially be overcome by assay-individualized translations of quantities into categories, which is also necessary for a meaningful clinical application.

The comparability of assays for TNF inhibitor drug levels and ADAb has been addressed in several previous studies. Vande Castele et al. [68] compared results from three ELISAs for IFX levels and ADAb. They found acceptable correlations, but limited agreement between the ELISAs, and concluded that there was a need for a standardization of drug monitoring assays. Steenholdt et al. [66] compared ELISA, radioimmunoassay and RGA for IFX levels and ADAb, and an enzyme immunoassay for IFX ADAb. They concluded that it was necessary to establish clinically relevant categories for each assay as a consequence of disagreement between assays for quantitative results.

Bloem et al. [61] compared several modalities to induce drug-tolerance in ADAb assays for the TNF inhibitor adalimumab. An acid-dissociation radioimmunoassay was found to be most sensitive for the detection of low levels of ADAb, resulting in a

large fraction of ADA-positive patients. The clinical significance of low titers of ADAb is, however, still unknown.

The presence of ADAb against the TNF inhibitor adalimumab at an early stage in the course of treatment has been suggested as a predictor for reduced response rates [72]. For some patients, low ADAb-titers may indicate the beginning of a clinically relevant immunogenic response, whereas other patients might be permanent or transient producers of low amounts of ADAb without clinical relevance. It has also been proposed that most patients exposed to biologics such as TNF inhibitors will react with some ADAb production, especially in the first phase of treatment, with yet-unknown clinical consequences [95].

Our data demonstrate that assays for drug levels and ADAb should not be used interchangeably, and that direct comparisons of quantitative drug levels obtained with different assays should be avoided. We and others have shown that even though drug monitoring by means of drug titer and ADAb measurement are implemented in clinical routine, there are unresolved questions and issues related to them.

# 2. Methodological background for mass cytometry studies

We set up a methodological background for mass cytometry studies including the collection, handling and stimulation of cells, and panel design.

## 2.1. Collection and storage of cells

Both PBMC and PBL (including bone marrow cells) were collected and used for studies included in this thesis. The most important differences between PBMC and PBL are (1) the ability to interfere with cells after storing, which only is possible in “live” cryopreserved PBMC, and (2) the conservation of granulocyte populations,

which only is given for PBL due to the loss of granulocytes in density centrifugation in PBMC (**Table 3**). The loss of granulocytes can be an advantage since the presence of large amounts of granulocytes can stand in the way for analysis of rarer immune cell subsets. Granulocyte subsets can also be a source of non-specific staining artefacts due to their content of eosinophil granules [96]. On the other hand, granulocytes as an important part and interactor of the innate immune system will be missing in the analysis of PBMC. Limited data in a concept study have been pointing at granulocytes as a population of interest regarding TNF signaling, supported by other data showing that granulocytes express high levels of TNF receptors [78, 97].

| PBMC | PBL |
|---|---|
| Do not contain granulocyte populations | Do contain granulocyte populations |
| Cells are exposed to about 30 min of centrifugation and the process of cryopreservation | Whole blood can be fixed immediately and while still being on body temperature |
| Cells can be cultured under standardized conditions after storage/freezing | Any interference with cells, e.g. cytokine stimulation, has to happen before fixation and storage |
| Density gradient centrifugation removes red blood cells, granulocytes and platelets. Additional steps of cell lysis are not necessary | Before analysis, red blood cells have to be removed by lysis, either before or after storage. |
| A viability stain must be part of analysis to sort out dead or compromised cells | All cells can be considered “live cells” until the moment of fixation and a viability stain is not necessary, if cells haven’t been exposed to more extensive perturbation before fixation |

**Table 3: Comparison Peripheral Blood Mononuclear Cells (PBMC) and Peripheral Blood Cells (PBL).** Advantages and disadvantages to the different ways of cell collection and storage.

A cryopreserved cell can be considered chemically and biologically extremely stable. However, depending on the osmotic tolerance of a cell, the toxicity of cryoprotectants, chilling and intracellular ice-formation, the processes around cryopreservation expose immune cells for stress and interfere with physiological patterns [98], especially with regards to rapid and more short-lived signaling pathways. Head-to-head comparisons between PBMC and PBL from the same donor regarding phenotypes and functional markers in immune cells are currently lacking.

For PBMC, factors of major concern are bench-time before cryopreservation, method and protocol for cell isolation, and resting time after thawing. In a set of quality control experiments, we addressed these concerns, and adapted our protocols to reduce variation and optimize signaling responses.

Limited data is available for comparisons between different live cell isolation techniques (Ficoll, CPT™ and SepMate™), showing similar performances for all techniques with slight advantages for CPT™ and SepMate™ in one study [99, 100]. Immediate fixation of immune cells in their physiological environment including stable body temperature and serum factors as performed for PBL, will reflect best in-vivo conditions, especially with regards to faster processes such as intracellular signaling and epitope phosphorylation. Depending on the question and project at hand, careful considerations should be taken in selecting collection SOPs.

## 2.2. Panel titration

Mass cytometry allows for simultaneous measurement of more than 50 parameters per single cell, including extra- and intracellular markers, barcodes, normalization beads etc. [85, 86, 94]. Compared to a flow cytometer, a mass cytometer has a narrower dynamic range, and a certain amount of spillover between channels has to be accepted. Spillover compensation is routinely applied to raw data in flow cytometry. For mass cytometry, tools for spillover compensation have been developed and suggested, but are currently not part of routine practice [101, 102]. In flow cytometry, signaling intensities of fluorescent dyes can vary largely, and markers of expected low abundance/intensity can be strategically conjugated to dyes of higher signaling intensity to provide a sufficient resolution. In mass cytometry, signaling intensities between different metal isotopes are fairly equal and certain amounts of spillover may have to be accepted to achieve the necessary resolution for markers of low abundance [81]. Signal spillover together with sources of non-specific antibody binding to secondary and low-affinity epitopes have to be taken into account in panel design and titration in mass cytometry. Due to the complexity of mass

cytometry panels, panel design and titration demand a deep understanding and extensive work.

Insufficient panel design and titration (bad data in) can result in signals and populations that do not exist or oversee existing epitopes and populations (bad data out). In **Paper 2**, we presented a pragmatic hierarchical approach to panel design and titration based on PBL and BM stained with a backbone panel to resolve more common cell populations with the help of the clustering algorithm SPADE, and different sub-panels for the titration of additional antibodies for deeper phenotyping and signaling. The same approach can be adapted to other cell types, both in suspension and tissue.

## 3. Signaling signatures in RA patients

We explored immune cell subsets for signaling signatures in RA patients compared to healthy individuals, with a primary focus on TNF signaling. Both B cell and T cell compartments have been investigated extensively in low-dimensional studies on patients with different rheumatic diseases [103, 104]. However, only few high-dimensional flow and mass cytometry studies have been published on classification and stratification of rheumatic patients. In two mass cytometry studies, O'Gorman et al. addressed chemokine and cytokine signatures in systemic and juvenile systemic lupus erythematosus [105]. They found that both newly diagnosed SLE patients and healthy donors had specific chemokine signatures especially in $CD14_{high}$ monocytes, and identified patient specific cytokine signatures in pediatric SLE, which could be evoked in cells from healthy donors upon patient-plasma exposure and disrupted by JAK1/JAK2 inhibition [106]. In the field of rheumatoid arthritis, Galligan et al. explored signaling patterns in a bigger, but rather inhomogeneous cohort of RA-patients by phospho-flow and found elevated levels of several phosphorylated epitopes in CD4 T cells in RA patients compared to healthy donors [77]. In a proof-of-principle-study on a single RA patient and one healthy donor, Nair et al. demonstrated that a complex mass cytometry setup may distinguish between health

and disease and may be able to detect changes after TNF inhibitor treatment [78]. Due to the illustrational character of their study, differences in signaling were not quantified, but both p-p38 and p-NFkB were differentiating markers in several cell subsets.

Very recently, Wang et al. used mass cytometry and next-generation sequencing to explore the B cell compartment of 16 patients with rheumatoid arthritis [107]. They found significant differences in naïve and memory B cell subsets of RA patients compared to healthy donors, including higher levels of serum IgA and IgM, indicating that B cell tolerance is altered in patients with RA compared to healthy donors. Next-generation sequencing revealed an Ig class switch in ACPA-positive RA patients towards fewer $IgA^+$ B cells and significantly lower levels of somatic hypermutations in RA patients.

Leong et al. studied 20 patients with juvenile idiopathic arthritis (JIA), who were good responders to anti-TNF and taken off treatment for 8 months. They identified a $CD3^+CD4^+CD45RA^-TNF^+$ T cell subset, which was elevated prior to therapy withdrawal in those JIA patients who relapsed after therapy withdrawal [108]. In the same study, transcriptomic profiling showed persistent dysregulations in pathways for T cell receptor-, apoptosis-, TNF-, NFkB- and MAPK-signaling in both relapse and remission patients compared to healthy donors, suggesting that such profiles may be used for the prediction of treatment responses rather than for the prediction of relapse after treatment discontinuation.

## 3.1. Limitations

There are currently limitations regarding the size and complexity of reasonable and feasible mass cytometry experiments. To identify differences between “healthy” and “sick”, representative cohorts of both groups are required. However, the number of simultaneously applicable barcodes, parallel handling of all samples, read-out time on the mass cytometer, and analysis of multi-dimensional data on millions of events set

currently limits on cohort sizes. For future studies with more samples, it is important to assure that results are robust across different cytometry runs, e.g., through the use of reference samples.

The core experiment of this thesis was the first major mass cytometry experiment carried out in Bergen, and the applied panel reflects both limited availability of metal-conjugated antibodies, and limited knowledge and experience in our group at the time of panel design.

PBMC were chosen as reporters of "RA-specific patterns" due to the possibility of simultaneous and standardized cell culturing and cytokine intervention after storing. For the purpose of characterization of cell states including fast and sensitive signaling pathways, we would retrospectively prefer the use of PBL, possibly combined with a strategy to reduce the total amount of granulocytes. Solutions such as SmartTubes™ offer new ways in the standardized collection of PBL including cytokine stimulation by adding fixed amounts of lyophilized cytokine to the tubes. Such solutions were not available when we started sample collection for our biobank.

We have studied signaling patterns in RA patients and healthy donors. Therefore, statements on the specificity of observed "RA patterns" compared to other inflammatory disorders cannot be made. Other TNF driven diseases, e.g. the spondylarthritides (SpA) or inflammatory bowel diseases (IBD), would be most suited for a comparison with RA patients with regards to TNF signaling and treatment responses to TNF inhibitors. Both RA, SpA and IBD share the TNF-associated inflammatory outcome, while differing significantly in underlying pathophysiology.

# Conclusion and future perspective

Cells, their functional status and behavior, and their cellular and non-cellular surroundings play a critical role in health and disease. Single cells from cell suspensions or tissue have been of extensive scientific interest ever since the first microscopic evidence of cells by Robert Hooke in 1665 [109]. Cells interact through direct cell-to-cell-contact or indirectly through messengers such as cytokines, chemokines, hormones, growth factors and antibodies. It has been and still is a challenge to grasp the wider picture of *cytomes*, heterogenous sets of cells that are functionally connected, such as blood cells or cancer tissue, while preserving a single cell resolution. Mass cytometry is another step in the direction of high-dimensional understanding of cytomes, both in suspension and in tissues. The abilities of mass cytometry have been demonstrated in a number of key studies on cells in suspension [110, 111] and tissues [112].

The pathophysiology of RA as well as the story of TNF – from the understanding of the role of TNF health and disease to the development of TNF inhibiting treatment – illustrate how central cytomics are. Our understanding of the complexity of biology, however, is currently largely built on studies, which due to technical limitations have contributed bits and pieces rather than providing a greater picture of biological processes.

Here, we illustrate the clinical challenge of precise diagnosis and patient stratification in RA and demonstrate ways of meeting the challenge with mass cytometry and tools for the analysis of high-dimensional data. The results of this thesis show that newly diagnosed RA patients can be classified correctly with relatively high precision based on signaling patterns in single cells, when compared to healthy donors. However, this thesis cannot offer a "comprehensive understanding of RA or TNF signaling" for the above-mentioned reasons related to material, cohorts and technicalities.

Other diseases with heavy involvement of TNF and standard application of TNF inhibitors are diseases from the family of spondylarthritides (e.g. psoriatic arthritis, ankylosing spondylitis) and from the family of inflammatory bowel diseases (Crohn's

disease, ulcerative colitis). We do not have material from newly diagnosed patients with these diagnoses at our disposal. However, we did collect an extensive material of PBMC from both RA and spondylarthritis (both psoriatic arthritis and ankylosing spondylitis) patients before and after initiation of TNF inhibitor treatment, which may enable us to explore signaling differences between cohorts as well as signaling differences between responders and non-responders to TNF inhibiting treatment.

The NordStar study, a recent international multicenter study on RA and patient stratification [113, 114], has provided us with the possibility of sample collection in patients randomized to treatment with either synthetic DMARD, TNF inhibitor, IL-6 inhibitor or CD80/86 co-stimulation inhibitor. Here, we collected PBL in SmartTubes™ with an unstimulated, a TNF stimulated and an IL-6 stimulated tube per patient before and 3 months after treatment initiation, and additional samples in case of treatment failure due to non-response or adverse events. Knowledge gained from our own study (**paper 3**) and the aforementioned studies by Wang et al. and Leong et al. [107, 108] may contribute to set up an interesting mass cytometry panel covering aspects of TNF and IL-6 signaling as well as T and B cell alterations before and under ongoing treatment.

Immune cells from peripheral blood may be sufficient reporters of drug responses, and may give a certain insight in the pathophysiology of RA. However, the primary site of inflammation in RA are joints and synovial tissue. We have started to collect cells from synovial fluid from RA patients in parallel with PBL from the same patients in a more recent project. Currently, no tissue mass cytometry studies have been undertaken on synovium sections, but a few studies on dissociated cells from synovial tissue [115-117]. After the advent of imaging mass cytometry [112] and its commercialization, analysis of synovial tissue sections on an image mass cytometer would be another interesting high-dimensional, single cell approach to the pathophysiology of RA. However, a standardized collection of synovial tissue from a well-sorted cohort of e.g. newly diagnosed RA patients has substantial practical and ethical implications due to the primary affection of small joints and the invasiveness of the procedure for sample taking.

We have made big efforts in standardized sample collection, established SOPs for all steps in complex mass cytometry experiments and gained experience in the handling and analysis of high-dimensional data. With this we hope to be able to contribute further to knowledge and future clinical applications.

# References


1. Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. The Lancet. 2009;373(9664):659-72.
2. Scott DL, Wolfe F, Huizinga TWJ. Rheumatoid arthritis. The Lancet. 2010;376(9746):1094-108.
3. Ahmed MS, Bae YS. Dendritic Cell-based Immunotherapy for Rheumatoid Arthritis: from Bench to Bedside. Immune Netw. 2016;16(1):44-51.
4. Smolen JS, Aletaha D, Barton A, Burmester GR, Emery P, Firestein GS, et al. Rheumatoid arthritis. Nat Rev Dis Primers. 2018;4:18001.
5. Nauts HC, Swift WE, Coley BL. The treatment of malignant tumors by bacterial toxins as developed by the late William B. Coley, M.D., reviewed in the light of modern research. Cancer Res. 1946;6:205-16.
6. Thomas L. The physiological disturbances produced by endotoxins. Annu Rev Physiol. 1954;16:467-90.
7. Carswell EA, Old LJ, Kassel RL, Green S, Fiore N, Williamson B. An endotoxin-induced serum factor that causes necrosis of tumors. Proc Natl Acad Sci U S A. 1975;72(9):3666-70.
8. Clark IA. How TNF was recognized as a key mechanism of disease. Cytokine and Growth Factor Reviews. 2007;18(3):335-43.
9. Aggarwal BB, Kohr WJ, Hass PE, Moffat B, Spencer SA, Henzel WJ, et al. Human tumor necrosis factor. Production, purification, and characterization. J Biol Chem. 1985;260(4):2345-54.
10. Beutler B, Greenwald D, Hulmes JD, Chang M, Pan YC, Mathison J, et al. Identity of tumour necrosis factor and the macrophage-secreted factor cachectin. Nature. 1985;316(6028):552-4.
11. Beutler B, Mahoney J, Le Trang N, Pekala P, Cerami A. Purification of cachectin, a lipoprotein lipase-suppressing hormone secreted by endotoxin-induced RAW 264.7 cells. J Exp Med. 1985;161(5):984-95.
12. Old LJ. Tumor necrosis factor (TNF). Science. 1985;230(4726):630-2.
13. Nedwin GE, Naylor SL, Sakaguchi AY, Smith D, Jarrett-Nedwin J, Pennica D, et al. Human lymphotoxin and tumor necrosis factor genes: structure, homology and chromosomal localization. Nucleic Acids Res. 1985;13(17):6361-73.
14. Pennica D, Nedwin GE, Hayflick JS, Seeburg PH, Derynck R, Palladino MA, et al. Human tumour necrosis factor: precursor structure, expression and homology to lymphotoxin. Nature. 1984;312(5996):724-9.
15. Mullberg J, Althoff K, Jostock T, Rose-John S. The importance of shedding of membrane proteins for cytokine biology. Eur Cytokine Netw. 2000;11(1):27-38.
16. Eissner G, Kolch W, Scheurich P. Ligands working as receptors: reverse signaling by members of the TNF superfamily enhance the plasticity of the immune system. Cytokine Growth Factor Rev. 2004;15(5):353-66.
17. Grell M, Douni E, Wajant H, Löhden M, Clauss M, Maxeiner B, et al. The transmembrane form of tumor necrosis factor is the prime activating ligand of the 80 kDa tumor necrosis factor receptor. Cell. 1995;83(5):793-802.
18. Aggarwal BB, Gupta SC, Kim JH. Historical perspectives on tumor necrosis factor and its superfamily: 25 years later, a golden journey. Blood. 2012;119(3):651-65.
19. Hehlgans T, Pfeffer K. The intriguing biology of the tumour necrosis factor/tumour necrosis factor receptor superfamily: players, rules and the games. Immunology. 2005;115(1):1-20.

20. Wajant H, Pfizenmaier K, Scheurich P. Tumor necrosis factor signaling. Cell Death Differ. 2003;10(1):45-65.
21. Pinckard JK, Sheehan KC, Schreiber RD. Ligand-induced formation of p55 and p75 tumor necrosis factor receptor heterocomplexes on intact cells. J Biol Chem. 1997;272(16):10784-9.
22. Brenner D, Blaser H, Mak TW. Regulation of tumour necrosis factor signalling: live or let die. Nat Rev Immunol. 2015;15(6):362-74.
23. Workman LM, Habelhah H. TNFR1 signaling kinetics: spatiotemporal control of three phases of IKK activation by posttranslational modification. Cell Signal. 2013;25(8):1654-64.
24. Gustin JA, Ozes ON, Akca H, Pincheira R, Mayo LD, Li Q, et al. Cell type-specific expression of the IkappaB kinases determines the significance of phosphatidylinositol 3-kinase/Akt signaling to NF-kappa B activation. J Biol Chem. 2004;279(3):1615-20.
25. Dayer JM, Beutler B, Cerami A. Cachectin/tumor necrosis factor stimulates collagenase and prostaglandin E2 production by human synovial cells and dermal fibroblasts. J Exp Med. 1985;162(6):2163-8.
26. Bertolini DR, Nedwin GE, Bringman TS, Smith DD, Mundy GR. Stimulation of bone resorption and inhibition of bone formation in vitro by human tumour necrosis factors. Nature. 1986;319(6053):516-8.
27. Di Giovine FS, Nuki G, Duff GW. Tumour necrosis factor in synovial exudates. Ann Rheum Dis. 1988;47(9):768-72.
28. Hopkins SJ, Meager A. Cytokines in synovial fluid: II. The presence of tumour necrosis factor and interferon. Clin Exp Immunol. 1988;73(1):88-92.
29. Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature. 1975;256(5517):495-7.
30. Knight DM, Trinh H, Le J, Siegel S, Shealy D, McDonough M, et al. Construction and initial characterization of a mouse-human chimeric anti-TNF antibody. Mol Immunol. 1993;30(16):1443-53.
31. Pedersen M. Kostnader til legemidler i helseforetak. Helsedirektoratet. 2018;1:24-5.
32. Kavanaugh A. Health economics: implications for novel antirheumatic therapies. Ann Rheum Dis. 2005;64 Suppl 4:iv65-9.
33. Kvamme MK, Lie E, Uhlig T, Moger TA, Kvien TK, Kristiansen IS. Cost-effectiveness of TNF inhibitors vs synthetic disease-modifying antirheumatic drugs in patients with rheumatoid arthritis: a Markov model study based on two longitudinal observational studies. Rheumatology (Oxford). 2015;54(7):1226-35.
34. Doan QV, Chiou CF, Dubois RW. Review of eight pharmacoeconomic studies of the value of biologic DMARDs (adalimumab, etanercept, and infliximab) in the management of rheumatoid arthritis. J Manag Care Pharm. 2006;12(7):555-69.
35. Pichler WJ. Adverse side-effects to biological agents. Allergy. 2006;61(8):912-20.
36. Bongartz T, Sutton AJ, Sweeting MJ, Buchan I, Matteson EL, Montori V. Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. Jama. 2006;295(19):2275-85.
37. Gartlehner G, Hansen RA, Jonas BL, Thieda P, Lohr KN. The comparative efficacy and safety of biologics for the treatment of rheumatoid arthritis: a systematic review and metaanalysis. J Rheumatol. 2006;33(12):2398-408.
38. Feuerstein JD, Cheifetz AS. Miscellaneous adverse events with biologic agents (excludes infection and malignancy). Gastroenterol Clin North Am. 2014;43(3):543-63.

39. Hernandez MV, Sanmarti R, Canete JD. The safety of tumor necrosis factor-alpha inhibitors in the treatment of rheumatoid arthritis. Expert Opin Drug Saf. 2016;15(5):613-24.
40. Cuppen BV, Welsing PM, Sprengers JJ, Bijlsma JW, Marijnissen AC, van Laar JM, et al. Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability. Rheumatology (Oxford). 2016;55(5):826-39.
41. Hueber W, Tomooka BH, Batliwalla F, Li W, Monach PA, Tibshirani RJ, et al. Blood autoantibody and cytokine profiles predict response to anti-tumor necrosis factor therapy in rheumatoid arthritis. Arthritis Res Ther. 2009;11(3):R76.
42. Pavy S, Toonen EJ, Miceli-Richard C, Barrera P, van Riel PL, Criswell LA, et al. Tumour necrosis factor alpha -308G->A polymorphism is not associated with response to TNFalpha blockers in Caucasian patients with rheumatoid arthritis: systematic review and meta-analysis. Ann Rheum Dis. 2010;69(6):1022-8.
43. Potter C, Hyrich KL, Tracey A, Lunt M, Plant D, Symmons DP, et al. Association of rheumatoid factor and anti-cyclic citrullinated peptide positivity, but not carriage of shared epitope or PTPN22 susceptibility variants, with anti-tumour necrosis factor response in rheumatoid arthritis. Ann Rheum Dis. 2009;68(1):69-74.
44. Thomas SS, Borazan N, Barroso N, Duan L, Taroumian S, Kretzmann B, et al. Comparative Immunogenicity of TNF Inhibitors: Impact on Clinical Efficacy and Tolerability in the Management of Autoimmune Diseases. A Systematic Review and Meta-Analysis. BioDrugs. 2015.
45. Rein P, Mueller RB. Treatment with Biologicals in Rheumatoid Arthritis: An Overview. Rheumatol Ther. 2017;4(2):247-61.
46. Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez-Olivo MA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. Cmaj. 2009;181(11):787-96.
47. Bendtzen K. Personalized medicine: theranostics (therapeutics diagnostics) essential for rational use of tumor necrosis factor-alpha antagonists. Discov Med. 2013;15(83):201-11.
48. Schaeverbeke T, Truchetet ME, Kostine M, Barnetche T, Bannwarth B, Richez C. Immunogenicity of biologic agents in rheumatoid arthritis patients: lessons for clinical practice. Rheumatology (Oxford). 2016;55(2):210-20.
49. Sethu S, Govindappa K, Alhaidari M, Pirmohamed M, Park K, Sathish J. Immunogenicity to Biologics: Mechanisms, Prediction and Reduction. Archivum Immunologiae et Therapiae Experimentalis. 2012;60(5):331-44.
50. Casadevall N, Nataf J, Viron B, Kolta A, Kiladjian JJ, Martin-Dupont P, et al. Pure red-cell aplasia and antierythropoietin antibodies in patients treated with recombinant erythropoietin. N Engl J Med. 2002;346(7):469-75.
51. Interferon beta-1b in the treatment of multiple sclerosis: final outcome of the randomized controlled trial. The IFNB Multiple Sclerosis Study Group and The University of British Columbia MS/MRI Analysis Group. Neurology. 1995;45(7):1277-85.
52. Elliott MJ, Maini RN, Feldmann M, Long-Fox A, Charles P, Bijl H, et al. Repeated therapy with monoclonal antibody to tumour necrosis factor alpha (cA2) in patients with rheumatoid arthritis. Lancet. 1994;344(8930):1125-7.
53. Maini RN, Breedveld FC, Kalden JR, Smolen JS, Davis D, Macfarlane JD, et al. Therapeutic efficacy of multiple intravenous infusions of anti-tumor necrosis factor alpha monoclonal antibody combined with low-dose weekly methotrexate in rheumatoid arthritis. Arthritis Rheum. 1998;41(9):1552-63.

54. Fogdell-Hahn A. Antidrug Antibodies: B Cell Immunity Against Therapy. Scand J Immunol. 2015;82(3):184-90.
55. Schultz HS, Reedtz-Runge SL, Backstrom BT, Lamberth K, Pedersen CR, Kvarnhammar AM. Quantitative analysis of the CD4+ T cell response to therapeutic antibodies in healthy donors using a novel T cell:PBMC assay. PLoS One. 2017;12(5):e0178544.
56. Boehncke WH, Brembilla NC. Immunogenicity of biologic therapies: causes and consequences. Expert Rev Clin Immunol. 2018:1-11.
57. Krieckaert CL, Bartelds GM, Lems WF, Wolbink GJ. The effect of immunomodulators on the immunogenicity of TNF-blocking therapeutic monoclonal antibodies: a review. Arthritis Res Ther. 2010;12(5):217.
58. Krieckaert CL, Nurmohamed MT, Wolbink GJ. Methotrexate reduces immunogenicity in adalimumab treated rheumatoid arthritis patients in a dose dependent manner. Ann Rheum Dis. 2012;71(11):1914-5.
59. Martin-Mola E, Balsa A, Garcia-Vicuna R, Gomez-Reino J, Gonzalez-Gay MA, Sanmarti R, et al. Anti-citrullinated peptide antibodies and their value for predicting responses to biologic agents: a review. Rheumatol Int. 2016;36(8):1043-63.
60. Lallemand C, Kavrochorianou N, Steenholdt C, Bendtzen K, Ainsworth MA, Meritet JF, et al. Reporter gene assay for the quantification of the activity and neutralizing antibody response to TNFalpha antagonists. J Immunol Methods. 2011;373(1-2):229-39.
61. Bloem K, van Leeuwen A, Verbeek G, Nurmohamed MT, Wolbink GJ, van der Kleij D, et al. Systematic comparison of drug-tolerant assays for anti-drug antibodies in a cohort of adalimumab-treated rheumatoid arthritis patients. J Immunol Methods. 2015;418:29-38.
62. Bendtzen K, Ainsworth M, Steenholdt C, Thomsen OO, Brynskov J. Individual medicine in inflammatory bowel disease: monitoring bioavailability, pharmacokinetics and immunogenicity of anti-tumour necrosis factor-alpha antibodies. Scand J Gastroenterol. 2009;44(7):774-81.
63. Gupta S, Indelicato SR, Jethwa V, Kawabata T, Kelley M, Mire-Sluis AR, et al. Recommendations for the design, optimization, and qualification of cell-based assays used for the detection of neutralizing antibody responses elicited to biological therapeutics. J Immunol Methods. 2007;321(1-2):1-18.
64. Lallemand C, Meritet J-F, Blanchard B, Lebon P, Tovey MG. One-step assay for quantification of neutralizing antibodies to biopharmaceuticals. Journal of Immunological Methods. 2010;356(1–2):18-28.
65. Bader LI, Solberg SM, Kaada SH, Bolstad N, Warren DJ, Gavasso S, et al. Assays for Infliximab Drug Levels and Antibodies: A Matter of Scales and Categories. Scand J Immunol. 2017;86(3):165-70.
66. Steenholdt C, Ainsworth MA, Tovey M, Klausen TW, Thomsen OO, Brynskov J, et al. Comparison of techniques for monitoring infliximab and antibodies against infliximab in Crohn's disease. Ther Drug Monit. 2013;35(4):530-8.
67. Steenholdt C, Bendtzen K, Brynskov J, Thomsen OO, Ainsworth MA. Clinical implications of measuring drug and anti-drug antibodies by different assays when optimizing infliximab treatment failure in Crohn's disease: post hoc analysis of a randomized controlled trial. Am J Gastroenterol. 2014;109(7):1055-64.
68. Vande Casteele N, Buurman DJ, Sturkenboom MG, Kleibeuker JH, Vermeire S, Rispens T, et al. Detection of infliximab levels and anti-infliximab antibodies: a comparison of three different assays. Aliment Pharmacol Ther. 2012;36(8):765-71.

69. Baert F, Noman M, Vermeire S, Van Assche G, D' Haens G, Carbonez A, et al. Influence of immunogenicity on the long-term efficacy of infliximab in Crohn's disease. N Engl J Med. 2003;348(7):601-8.
70. Targan SR, Hanauer SB, van Deventer SJ, Mayer L, Present DH, Braakman T, et al. A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factor alpha for Crohn's disease. Crohn's Disease cA2 Study Group. N Engl J Med. 1997;337(15):1029-35.
71. Suzuki T, Tada M, Ishii-Watabe A. Development of anti-drug monoclonal antibody panels against adalimumab and infliximab. Biologicals. 2019.
72. Jani M, Chinoy H, Warren RB, Griffiths CE, Plant D, Fu B, et al. Clinical utility of random anti-tumor necrosis factor drug-level testing and measurement of antidrug antibodies on the long-term treatment response in rheumatoid arthritis. Arthritis Rheumatol. 2015;67(8):2011-9.
73. Jani M, Gavan S, Chinoy H, Dixon WG, Harrison B, Moran A, et al. A microcosting study of immunogenicity and tumour necrosis factor alpha inhibitor drug level tests for therapeutic drug monitoring in clinical practice. Rheumatology (Oxford). 2016.
74. Krieckaert CL, Nair SC, Nurmohamed MT, van Dongen CJ, Lems WF, Lafeber FP, et al. Personalised treatment using serum drug levels of adalimumab in patients with rheumatoid arthritis: an evaluation of costs and effects. Ann Rheum Dis. 2015;74(2):361-8.
75. Steenholdt C, Brynskov J, Thomsen OO, Munck LK, Fallingborg J, Christensen LA, et al. Individualised therapy is more cost-effective than dose intensification in patients with Crohn's disease who lose response to anti-TNF treatment: a randomised, controlled trial. Gut. 2013.
76. Ermann J, Rao DA, Teslovich NC, Brenner MB, Raychaudhuri S. Immune cell profiling to guide therapeutic decisions in rheumatic diseases. Nat Rev Rheumatol. 2015.
77. Galligan CL, Siebert JC, Siminovitch KA, Keystone EC, Bykerk V, Perez OD, et al. Multiparameter phospho-flow analysis of lymphocytes in early rheumatoid arthritis: implications for diagnosis and monitoring drug therapy. PLoS One. 2009;4(8):e6703.
78. Nair N, Mei HE, Chen SY, Hale M, Nolan GP, Maecker HT, et al. Mass cytometry as a platform for the discovery of cellular biomarkers to guide effective rheumatic disease therapy. Arthritis Res Ther. 17. England2015. p. 127.
79. Bandura DR, Baranov VI, Ornatsky OI, Antonov A, Kinach R, Lou X, et al. Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. Anal Chem. 2009;81(16):6813-22.
80. Ornatsky O, Bandura D, Baranov V, Nitz M, Winnik MA, Tanner S. Highly multiparametric analysis by mass cytometry. J Immunol Methods. 2010;361(1-2):1-20.
81. Ornatsky O, Baranov VI, Bandura DR, Tanner SD, Dick J. Multiple cellular antigen detection by ICP-MS. J Immunol Methods. 2006;308(1-2):68-76.
82. Chattopadhyay PK, Roederer M. A mine is a terrible thing to waste: high content, single cell technologies for comprehensive immune analysis. Am J Transplant. 2015;15(5):1155-61.
83. Bendall SC, Nolan GP, Roederer M, Chattopadhyay PK. A deep profiler's guide to cytometry. Trends Immunol. 2012;33(7):323-32.
84. Finck R, Simonds EF, Jager A, Krishnaswamy S, Sachs K, Fantl W, et al. Normalization of mass cytometry data with bead standards. Cytometry A. 2013;83(5):483-94.

85. Zunder ER, Finck R, Behbehani GK, Amir el AD, Krishnaswamy S, Gonzalez VD, et al. Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. Nat Protoc. 2015;10(2):316-33.
86. Kleinsteuber K, Corleis B, Rashidi N, Nchinda N, Lisanti A, Cho JL, et al. Standardization and quality control for high-dimensional mass cytometry studies of human samples. Cytometry A. 2016;89(10):903-13.
87. Qiu P, Simonds EF, Bendall SC, Gibbs KD, Jr., Bruggner RV, Linderman MD, et al. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nat Biotechnol. 2011;29(10):886-91.
88. Amir el AD, Davis KL, Tadmor MD, Simonds EF, Levine JH, Bendall SC, et al. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat Biotechnol. 2013;31(6):545-52.
89. Bruggner RV, Bodenmiller B, Dill DL, Tibshirani RJ, Nolan GP. Automated identification of stratifying signatures in cellular subpopulations. Proc Natl Acad Sci U S A. 2014;111(26):E2770-7.
90. Bagwell B. A New Analytic Approach for Live Singlet Identification 2017 [updated 09.06.2017. Available from: https://www.fluidigm.com/articles/mass-cyto-summit-videos?mkt_tok=eyJpIjoiT0RVeU5ERTBNbU5oWW1GbSIsInQiOiJ4U3NDdUZjbVoyRCt6UThrWXg4b3M3RUNick5ZK0lsekJZQzU3VFZyU0QwYW5UaHdaQUxwbDh3ZHJzbTRcLzg3MlhIWkh5bEFDMUdWSFk0NWpiTmhueGVXQ3IzQzI4K0Q5a015cGM1bFdrVHl2elpNY2JGUERqb2I3MEk0RW9FemgifQ%3D%3D#bagwell.
91. Fluidigm. Approach to Bivariate Analysis of Data Acquired Using the Maxpar Human Immune Monitoring Panel Kit. 2018:4-7.
92. Qiu P. Toward deterministic and semiautomated SPADE analysis. Cytometry A. 2017;91(3):281-9.
93. Jorgensen KK, Olsen IC, Goll GL, Lorentzen M, Bolstad N, Haavardsholm EA, et al. Switching from originator infliximab to biosimilar CT-P13 compared with maintained treatment with originator infliximab (NOR-SWITCH): a 52-week, randomised, double-blind, non-inferiority trial. Lancet. 2017;389(10086):2304-16.
94. Spitzer MH, Nolan GP. Mass Cytometry: Single Cells, Many Features. Cell. 2016;165(4):780-91.
95. van Schouwenburg PA, Krieckaert CL, Rispens T, Aarden L, Wolbink GJ, Wouters D. Long-term measurement of anti-adalimumab using pH-shift-anti-idiotype antigen binding test shows predictive value and transient antibody formation. Ann Rheum Dis. 2013;72(10):1680-6.
96. Rahman AH, Tordesillas L, Berin MC. Heparin reduces nonspecific eosinophil staining artifacts in mass cytometry experiments. Cytometry A. 2016;89(6):601-7.
97. Liu C, Tang J. Expression levels of tumor necrosis factor-alpha and the corresponding receptors are correlated with trauma severity. Oncol Lett. 2014;8(6):2747-51.
98. Woods EJ, Thirumala S, Badhe-Buchanan SS, Clarke D, Mathew AJ. Off the shelf cellular therapeutics: Factors to consider during cryopreservation and storage of human cells for clinical use. Cytotherapy. 2016;18(6):697-711.
99. Corkum CP, Ings DP, Burgess C, Karwowska S, Kroll W, Michalak TI. Immune cell subsets and their gene expression profiles from human PBMC isolated by Vacutainer Cell Preparation Tube (CPT) and standard density gradient. BMC Immunol. 2015;16:48.
100. Grievink HW, Luisman T, Kluft C, Moerland M, Malone KE. Comparison of Three Isolation Techniques for Human Peripheral Blood Mononuclear Cells: Cell Recovery

and Viability, Population Composition, and Cell Functionality. Biopreserv Biobank. 2016;14(5):410-5.
101. Chevrier S, Crowell HL, Zanotelli VRT, Engler S, Robinson MD, Bodenmiller B. Compensation of Signal Spillover in Suspension and Imaging Mass Cytometry. Cell Syst. 2018;6(5):612-20.e5.
102. Sekhri P, Kim MY, Behbehani GK. Unlabeled Competitor Antibody to Reduce Nonlinear Signal Spillover in Mass Cytometry. Cytometry A. 2019.
103. Wei C, Jenks S, Sanz I. Polychromatic flow cytometry in evaluating rheumatic disease patients. Arthritis Res Ther. 2015;17:46.
104. Fonseka CY, Rao DA, Raychaudhuri S. Leveraging blood and tissue CD4+ T cell heterogeneity at the single cell level to identify mechanisms of disease in rheumatoid arthritis. Curr Opin Immunol. 2017;49:27-36.
105. O'Gorman WE, Hsieh EWY, Savig ES, Gherardini PF, Hernandez JD, Hansmann L, et al. Single-cell systems-level analysis of human Toll-like receptor activation defines a chemokine signature in patients with systemic lupus erythematosus. Journal of Allergy and Clinical Immunology. 2015;136(5):1326-36.
106. O'Gorman WE, Kong DS, Balboni IM, Rudra P, Bolen CR, Ghosh D, et al. Mass cytometry identifies a distinct monocyte cytokine signature shared by clinically heterogeneous pediatric SLE patients. J Autoimmun. 2017.
107. Wang Y, Lloyd KA, Melas I, Zhou D, Thyagarajan R, Lindqvist J, et al. Rheumatoid arthritis patients display B-cell dysregulation already in the naive repertoire consistent with defects in B-cell tolerance. Sci Rep. 2019;9(1):19995.
108. Leong JY, Chen P, Yeo JG, Ally F, Chua C, Nur Hazirah S, et al. Immunome perturbation is present in patients with juvenile idiopathic arthritis who are in remission and will relapse upon anti-TNFalpha withdrawal. Ann Rheum Dis. 2019.
109. Gest H. The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. Notes Rec R Soc Lond. 2004;58(2):187-201.
110. Bendall SC, Simonds EF, Qiu P, Amir el AD, Krutzik PO, Finck R, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science. 2011;332(6030):687-96.
111. Fragiadakis GK, Gaudilliere B, Ganio EA, Aghaeepour N, Tingle M, Nolan GP, et al. Patient-specific Immune States before Surgery Are Strong Correlates of Surgical Recovery. Anesthesiology. 2015;123(6):1241-55.
112. Giesen C, Wang HA, Schapiro D, Zivanovic N, Jacobs A, Hattendorf B, et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nat Methods. 2014;11(4):417-22.
113. Lund Hetland M, Haavardsholm E, Rudin A, Nordström D, Nurmohamed M, Gudbjornsson B, et al. A Multicenter Randomized Study in Early Rheumatoid Arthritis to Compare Active Conventional Therapy versus Three Biological Treatments: 24 Week Efficacy and Safety Results of the NORD-STAR Trial [abstract]. Arthritis Rheumatol. 2019;71.
114. Glinatsi D, Heiberg MS, Rudin A, Nordstrom D, Haavardsholm EA, Gudbjornsson B, et al. Head-to-head comparison of aggressive conventional therapy and three biological treatments and comparison of two de-escalation strategies in patients who respond to treatment: study protocol for a multicenter, randomized, open-label, blinded-assessor, phase 4 study. Trials. 2017;18(1):161.
115. Donlin LT, Rao DA, Wei K, Slowikowski K, McGeachy MJ, Turner JD, et al. Methods for high-dimensional analysis of cells dissociated from cryopreserved synovial tissue. Arthritis Res Ther. 2018;20(1):139.

116. Fonseka CY, Rao DA, Teslovich NC, Korsunsky I, Hannes SK, Slowikowski K, et al. Mixed-effects association of single cells identifies an expanded effector CD4(+) T cell subset in rheumatoid arthritis. Sci Transl Med. 2018;10(463).
117. Zhang F, Wei K, Slowikowski K, Fonseka CY, Rao DA, Kelly S, et al. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. Nat Immunol. 2019;20(7):928-42.

# Papers 1-3

**Paper 1:** Assays for Infliximab Drug Levels and Antibodies: A Matter of Scales and Categories.

**Paper 2:** Titrating Complex Mass Cytometry Panels.

Supplementary material:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6766997/bin/CYTO-95-792-s001.doc

**Paper 3:** Candidate markers for stratification and classification in rheumatoid arthritis

Supplementary material:
https://www.frontiersin.org/articles/10.3389/fimmu.2019.01488/full#supplementary-material

# I

doi: 10.1111/sji.12572

# Assays for Infliximab Drug Levels and Antibodies: A Matter of Scales and Categories

L. I. Bader*,†, S. M. Solberg‡,§, S. H. Kaada¶, N. Bolstad**, D. J. Warren**, S. Gavasso*,††, C. G. Gjesdal*,† & C. A. Vedeler††,‡‡

*The Bergen Group of Epidemiology and Biomarkers in Rheumatic Disease (BEaBiRD), Department of Rheumatology, Haukeland University Hospital, Bergen, Norway; †Department of Clinical Science, University of Bergen, Bergen, Norway; ‡Department of Dermatology, Haukeland University Hospital, Bergen, Norway; §Broegelmann Research Laboratory, Department of Clinical Science, University of Bergen, Bergen, Norway; ¶Department of Immunology and Transfusion Medicine, Haukeland University Hospital, Bergen, Norway; **Department of Medical Biochemistry, Oslo University Hospital, Radiumhospitalet, Oslo, Norway; ††Department of Clinical Medicine, University of Bergen, Bergen, Norway; and ‡‡Department of Neurology, Haukeland University Hospital, Bergen, Norway

*Received 18 April 2017; Accepted in revised form 23 May 2017*

*Correspondence to*: C. Vedeler, Department of Neurology, Haukeland University Hospital, Post Office Box 1400, 5021 Bergen, Norway. E-mail: christian.vedeler@helse-bergen.no

## Abstract

Immunogenicity is a frequent cause of secondary non-response to tumour necrosis factor (TNF) inhibitors. Drug level measurement and detection of antidrug antibodies have been shown to be cost effective and clinically relevant, and a large number of assays are available for these purposes. It is, however, difficult to compare assays and translate results into clinical meaningful information due to different methodological approaches and a lack of assay standardization. We have analysed infliximab drug levels and antidrug antibodies in 107 patient samples using enzyme-linked immunoassays (ELISA), immunofluorometric assays (IFMA) and reporter-gene assays (RGA). The RGA gave the lowest results for drug levels, whereas the IFMA detected the highest number of antidrug antibody positive sera. Applying individualized therapeutic ranges to each assay resulted in agreement among all three assays in 74% of samples for drug levels and 98% of samples for antidrug antibodies. We found that TNF inhibitor monitoring assays measure on different scales and that the agreement between quantitative results is limited. However, interassay differences can partially be overcome by assay-individualized translations of quantities into categories, which also is necessary for a meaningful clinical application. Our data demonstrate that assays should not be used interchangeably and that direct comparison of quantitative drug levels obtained with different assays should be avoided.

## Introduction

Tumour necrosis factor (TNF) inhibitors are widely used for treatment of inflammatory joint, skin and bowel diseases [1–3]. TNF inhibitors have immunogenic properties that can be the cause of secondary non-response and injection- or infusion-related side effects [4–6]. The clinical implications of TNF inhibitor immunogenicity are complex due to disease heterogeneity, different treatment and dosing regimens and co-medications. Furthermore, not all cases of secondary non-response are due to the immunogenic properties of the drug [7].

A large number of assays are available to identify patients with immunogenic reactions to TNF inhibitors, usually by means of drug level measurement and detection of antidrug antibodies (ADA) [7–10]. TNF inhibitor drug monitoring has been shown to be cost effective and clinically relevant [9, 11]. In a recent meta-analysis covering over 14,000 patients, ADAs were present in 12.7% of patients treated with a TNF inhibitor, and the presence of ADA was found to reduce the odds of drug response by 67% [12]. In the same study, the presence of ADA was shown to increase the risk of reaction at injection or infusion sites (OR 3.25), whereas co-medication with other immunosuppressants was shown to reduce the odds for ADA production by 74%.

TNF inhibitor drug monitoring assays employ a variety of methodological approaches. Assays for drug levels, as opposed to ADA assays, employ standardized units and can therefore be compared directly. Several studies have found acceptable interassay correlations. However, interassay agreement, that is, the true accordance between the results of two or more assays was less satisfying [13–15]. Clinicians and clinical laboratories are often provided with a number of cut-offs, ranges, limits and concentrations, but not with clinical reference ranges





SCANDINAVIAN JOURNAL OF Immunology AN INTERNATIONAL MONTHLY JOURNAL

for drug levels. Thus, clinicians have to determine themselves the drug levels to consider as 'therapeutic' or 'non-therapeutic'.

ADA assays, on the other hand, do not have standardized units as results are presented in arbitrary units (AU) or AU/mL. Therefore, a direct comparison between assays is only possible if ADA quantities are transformed into categories such as 'negative' or 'positive'. For ADA, both quantity and quality of the identified antibody depend on the applied assay methodology. The quantity of detected ADA is affected by the drug-tolerance of an assay. Many assays do not detect ADA bound to the drug; they are drug-sensitive. Some drug-tolerant assays overcome this problem, for example by dissociation of ADA from the drug. Assays also vary in their abilities to distinguish biologically functional from non-functional ADA. Functional assays detect only ADAs that neutralize the function of the drug (neutralizing antibodies, NAb). Non-functional assays detect ADA binding to the drug (binding antibodies, BAb) and, thus, detect antibodies that do not affect the drug functionally. Being solely confronted with a 'positive' result of an ADA assay, without information about quantity or quality of the identified ADA, the clinician will not be able to decide relevance for the patient.

In this study, we illustrate the issue of lacking scales and categories by comparing three assays for infliximab (IFX) drug levels and three assays for IFX ADA. Our data demonstrate that direct comparison of quantitative drug levels obtained with different assays should be avoided. We found, however, that interassay differences can partially be overcome by assay-individualized translations of quantities into categories.

## Materials and methods

*Patients and serum samples*. One hundred and seven patients treated with IFX at the Departments of Rheumatology ($n$ = 78) and Dermatology ($n$ = 29) at the Haukeland University Hospital, Bergen, Norway participated in this observational cross-sectional study. Serum samples were collected and analysed between April 2014 and May 2015. A serum sample from each patient was collected at a randomly selected infusion appointment just prior to the IFX infusion. All patients gave written informed consent for biobank inclusion at the Departments of Rheumatology and Dermatology (regional ethics committee approvals 2012/1689 and 2014/1373).

*Assays*. Methods, thresholds and ranges of assays for IFX levels and IFX ADA are given in Table 1. IFX serum levels were measured with a capture enzyme-linked immunosorbent assay (ELISA; Promonitor®-IFX), whereas a bridging ELISA (anti-IFX/Orion Diagnostica) was used for the detection and quantification of IFX ADA. Automated immunofluorometric assays (IFMAs) were performed on the AutoDELFIA® immunoassay platform. Briefly, the method measures drug levels using a target-based assay with TNF on solid phase; the active drug in the patient sample binds TNF and is detected with a europium-labelled tracer that binds the Fc-domain on the drug. Sera with IFX levels <10 mg/L were analysed for IFX ADA using an IFMA that detects only antibodies binding to the TNF-binding site of the drug. Infliximab and infliximab NAb reporter-gene assays (RGAs; iLite®/Biomonitor) are based on an immortalized and modified erythroleukemic cell line using the same biological principle for both IFX and ADA quantification. TNF binding to its cognate receptor at the cell

**Table 1** Methods, thresholds and ranges of assays for IFX levels and IFX ADA.

| | ELISA | IFMA | RGA |
|---|---|---|---|
| IFX levels | | | |
| Method | Capture ELISA | Target-based assay: TNF on solid phase, the drug-containing patient-serum in the liquid phase and a tracer binding the drug | Based on a reporter cell line: TNF signalling induces expression of firefly luciferase |
| Manufacturer provided categories and therapeutic ranges | <0.035 μg/mL: Negative<br>≥0.035 μg/mL: Positive<br>0.035-1.5 μg/mL: Subtherapeutic<br>>1.5 μg/mL: Therapeutic | <1 mg/L: Negative<br>≥1 mg/L: Positive<br>2–12 mg/L: Wide therapeutic range<br>3–8 mg/L: Narrow therapeutic range | <0.65 μg/mL: Negative<br>≥0.65 μg/mL: Positive<br>No therapeutic ranges provided |
| IFX ADA | | | |
| Method | Bridging ELISA | Immunofluorometric assay: detects only antibodies against the TNF-binding site of the drug | Based on a reporter cell line: TNF signalling induces expression of firefly luciferase |
| Manufacturer provided categories | ≤5 AU/L: Negative<br>>5 AU/L: Positive | <10 AU/L: Negative<br>10–30 AU/L: Low-positive<br>30–80 AU/L: Positive<br>>80 AU/L: High-positive | Positive and negative are defined by an internal threshold |

IFX, infliximab; ADA, antidrug antibodies; ELISA, enzyme-linked immunosorbent assay; IFMA, immunofluorometric assay; RGA, reporter-gene assay; AU, arbitrary units.

surface activates intracellularly the transcription factor NFκB, resulting in expression of firefly luciferase. TNF signalling activity is quantified by detection of firefly luciferase luminescence. The assay uses *Renilla* luminescence for normalization. Serum with IFX levels below the lowest reported concentration of 0.65 μg/mL was analysed further for the presence of IFX NAb. The reporter-gene assay is a drug-sensitive and functional assay [16].

*Data collection and statistical analysis*. All clinical data were obtained from the hospital files, the Norwegian Arthritis Registry (NorArtritt) and the local quality registry at the Department of Dermatology. Statistical analysis was performed using the statistics package SPSS (version 22, IBM). Agreement in quantitative results from the three assays for IFX levels was assessed and visualized in Bland–Altman plots [17]. In these plots, the average of results from all three assays for one patient sample is plotted on the *x*-axis against the difference between the two assays of interest on the *y*-axis. If two assays are in complete agreement, the data will fall along a horizontal line through zero, whereas disagreement will result in a distribution above or below that line. A deviation of the mean difference away from 0 that is not attributable to extreme outliers indicates that there is a systematic difference between assays. The effects of transformation from quantitative to qualitative and categorical data on interassay agreement are displayed in Venn diagrams. Agreement in quantitative results for IFX ADA could not be assessed due to arbitrary units for antibodies. Interassay agreement after transformation of quantitative data into the categories negative and positive is presented in Venn diagrams.

There are no numerical values for good or poor agreement; based on the underlying question the observer must decide whether the degree of agreement is satisfying or not.

## Results

### Patients and treatment

One hundred and seven patients (53 females and 54 males) were included; mean age and disease duration were 51 years and 19 years, respectively. Ninety-four patients were treated with IFX-Remicade® (Janssen Biologics), and 13 patients were treated with IFX-Remsima® (Celltrion). Patient characteristics, diagnoses, treatment and co-medication details are given in Table 2.

**Table 2** Patient characteristics, diagnoses, medication and co-medication.

| | |
|---|---|
| Number of patients | 107 |
| Female/male | 53/54 |
| Mean patient age (min-max) | 51 years (16–86) |
| Diagnoses, number of patients (% of total) | |
| Rheumatoid arthritis | 37 (34.6) |
| Spondyloarthritis (excl. psoriatic arthritis) | 18 (16.8) |
| Psoriatic arthritis | 17 (15.9) |
| Juvenile idiopathic arthritis | 9 (8.4) |
| Psoriasis | 23 (21.5) |
| Others | 3 (2.8) |
| Mean disease duration (min-max) | 19 years (2–50) |
| Medication, infliximab (IFX) | |
| IFX-Remicade (%) | 94 (88) |
| IFX-Remsima (%) | 13 (12) |
| Median treatment duration (min-max) | 70 months (0–158) |
| Mean IFX dose (min-max) | 4.5 mg/kg (2.1–10.4) |
| Mean IFX dose rheumatology (min-max) | 3.8 mg/kg (2.1–7.3) |
| Mean IFX dose dermatology (min-max) | 6.4 mg/kg (4.5–10.4) |
| Mean IFX interval (min-max) | 7.5 weeks (4–17) |
| Co-medication | |
| Methotrexate | 76 patients |
| Other disease-modifying antirheumatic drugs | 2 sulfasalazine, 1 leflunomide, 1 azathioprine |
| Prednisolone | 8 patients |

### Infliximab drug levels

The three assays were in better agreement at lower IFX levels and poorer for higher levels. No samples with an average IFX level <10 mg/L were outside the limits of agreement, which were narrowest for ELISA versus IFMA. The RGA resulted in systematically lower IFX levels than the ELISA, whereas the IFMA resulted in higher levels than the ELISA. Analysis of agreement among quantitative results from all three assays is presented in Figure 1A–C.

We evaluated several therapeutic drug ranges as shown in Figure 2A. Application of two fixed therapeutic ranges for all three assays (one narrower, one wider), resulted in low rates of agreement for about 50% of the patient samples, whereas an individual adaptation of categories to each assay improved the rate of agreement to 74% of all patient samples. For all categories, agreement rates were significantly higher between ELISA and IFMA than ELISA and RGA or IFMA and RGA.

### Infliximab antidrug antibodies (IFX ADA)

Of the 107 patient samples, four were IFX ADA positive in the ELISA (≥5 AU/mL). The same four samples were found to be positive for ADA by the RGA as was one additional sample that was just below the threshold in the ELISA (4.7 AU/mL). The IFMA identified a total of 17 ADA-positive samples, 11 of which fell into the category 'low positive' (10-30 AU/L). The remaining six samples included the ADA-positive samples from ELISA and RGA. One patient, who was classified as IFX ADA positive in all three assays, was treated with Remsima®.

IFX was detected in only one of the ADA-positive samples at a low concentration (1.7 mg/L) by the IFMA. All of the other ADA-positive samples had IFX levels below threshold in all assays. Agreement for categorical results for ADA from all three assays is presented in Figure 2B.

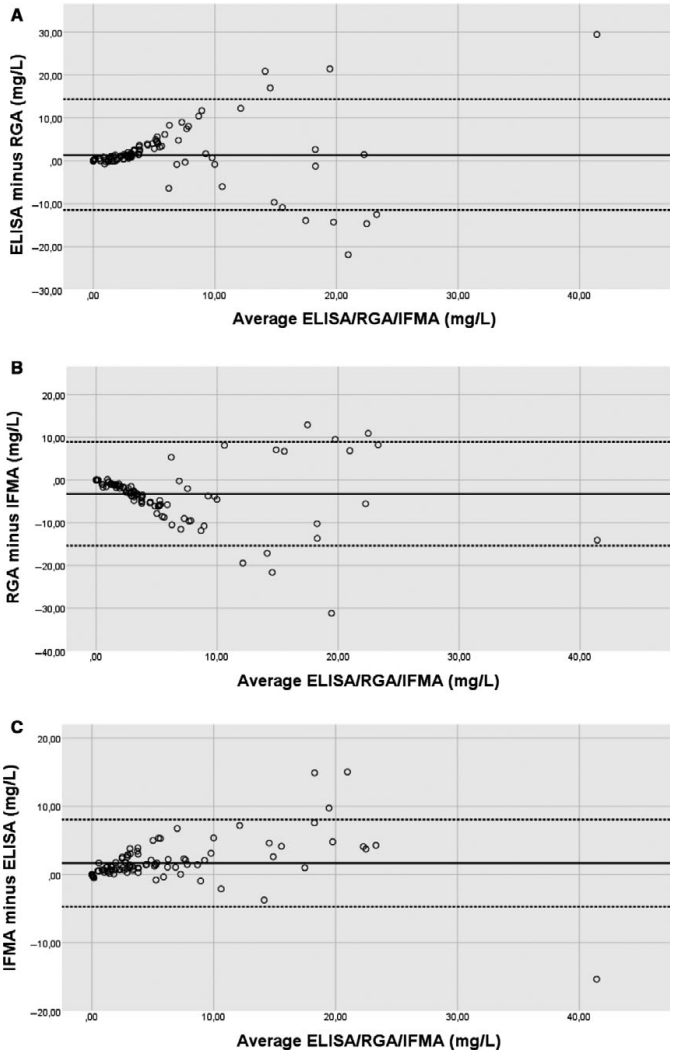



**Figure 1** Bland–Altman plots demonstrating agreement between IFX levels determined using different techniques. The average result from all three assays for one patient sample (*x*-axis) is plotted against the difference between the two assays compared (*y*-axis) for A: ELISA versus RGA, B: RGA versus IFMA, and C: IFMA versus ELISA. The difference is 0 in case of total agreement; increasing distance from 0 indicates an increasing disagreement between the two compared assays.

## Discussion

For IFX drug levels, we found agreement between ELISA and IFMA of 80% or higher, depending on the applied therapeutic range. Quantitatively, the IFMA resulted in higher drug levels, and the RGA resulted in lower drug levels. For IFX ADA, the ELISA classified four samples as ADA positive, RGA classified five as positive and the IFMA identified 17 positive samples. For both IFX levels and IFX ADA, adaptation of categories such as 'therapeutic' or 'positive' to each assay resulted in comparable performances of all assays with agreement rates of 74% (IFX levels) and 98% (IFX ADA), respectively. We did not include testing for repeatability as all assays undergo such evaluation in their validation process. Intra and between-day variations can influence the results especially for the detection of ADA, as was shown in an assay comparison reported by Steenholdt *et al.* [13].

In this study, we focused on assay methodology for clinical interpretation. For clinical decision-making, the individual clinical setting provides most relevant longitudinal information [7, 8, 18]. The issue of scales and

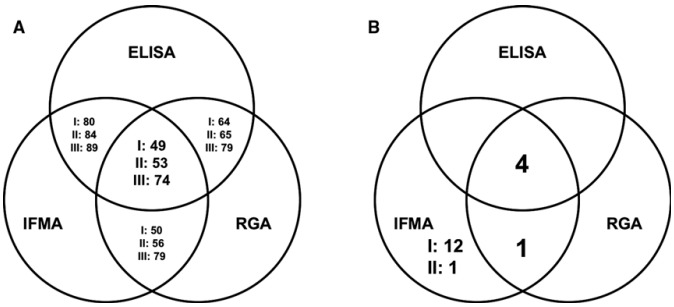



**Figure 2** Venn diagrams illustrating agreement for categorical results in ELISA, IFMA and RGA. (A) The percentages of samples with agreement on IFX levels in all three assays and between two assays. I: therapeutic range of 3–8 mg/L for all assays; II: therapeutic range of 1.5–12 mg/L for all assays; and III: data in individualized ranges of RGA 1–9 mg/L, ELISA 2-10 mg/L and IFMA 3–11 mg/L B: The number of ADA-positive patient samples (of the total of 107) are shown as follows: four samples were detected as positive all assays, 1 sample was positive in the RGA and the IFMA, 12 samples were positive only by IFMA. Removal of all 'low-positive' samples from IFMA resulted in one positive sample only detected by the IFMA and an overall agreement on the outcome of 98% of samples.

categories is not only important for the clinical application of assay results but also for the understanding of variations presented in studies on TNF inhibitor immunogenicity. Especially for the detection of ADA, there is a wide range of prevalence, which is related to the use of particular assays and interpretation of data. For example, the reported prevalence of IFX ADA was 5% in one study and 61% in another [4, 19].

The comparability of assays for TNF inhibitor drug levels and ADA has been addressed in several previous studies. Vande Castele *et al.*[15] compared results from three ELISAs for IFX levels and ADA. They found acceptable correlations, but limited agreement between the ELISAs, and concluded that there was a need for a standardization of drug monitoring assays. Steenholdt *et al.*[13] compared ELISA, radioimmunoassay and RGA for IFX levels and ADA, and an enzyme immunoassay for IFX ADA. They concluded that it was necessary to establish clinically relevant categories for each assay as a consequence of disagreement between assays for quantitative results.

Bloem *et al.*[20] compared several modalities to induce drug-tolerance in ADA assays for the TNF inhibitor adalimumab. An acid-dissociation radioimmunoassay was found to be most sensitive for the detection of low levels of ADA, resulting in a large fraction of ADA-positive patients. The clinical significance of low titres of ADA is, however, still unknown.

The presence of ADA against adalimumab at an early stage in the course of treatment has been suggested as a predictor for reduced response rates [21]. For some patients, low ADA-titres may indicate the beginning of a clinically relevant immunogenic response, whereas other patients might be permanent or transient producers of low amounts of ADA without clinical relevance. It has also been proposed that most patients exposed to biologics such as TNF inhibitors will react with some ADA production, especially in the first phase of treatment, with yet-unknown clinical consequences [22].

We found that TNF inhibitor monitoring assays measure on different scales and that the agreement between the assays is limited. However, interassay differences can be overcome by an assay-individualized translation of quantities into categories, which also are necessary for meaningful clinical application. In order to be able to handle the issue of scales and categories and to translate results into clinically meaningful information, manufacturers and clinical laboratories should not offer TNF inhibitor drug monitoring assays without providing guiding categories and information on the assay method, especially on drug-tolerance and assay functionality for ADA detection. ADA values could be categorized as negative, low level and high level and drug levels with a therapeutic range. Clinicians making use of these data should have a general understanding of the assay methods to be able to interpret and implement results. Furthermore, clinicians should avoid using assays interchangeably and should not attempt to directly compare quantitative drug levels between assays.

## Disclosure statements

iLite® kits (RGA) and luminometer used for this study were donated by Pfizer Norway to the Neuroimmunology Laboratory at Haukeland University Hospital without restrictions. Nils Bolstad has received honoraria from Pfizer (advisory board), Orion Pharma (advisory board, lecture) and Napp Pharmaceuticals (lecture).

## References

1 Knight DM, Trinh H, Le J *et al.* Construction and initial characterization of a mouse-human chimeric anti-TNF antibody. *Mol Immunol* 1993;30:1443–53.

2 Tracey D, Klareskog L, Sasso EH, Salfeld JG, Tak PP. Tumor necrosis factor antagonist mechanisms of action: a comprehensive review. *Pharmacol Ther* 2008;117:244–79.
3 Willrich MA, Murray DL, Snyder MR. Tumor necrosis factor inhibitors: clinical utility in autoimmune diseases. *Transl Res* 2015;165:270–82.
4 Baert F, Noman M, Vermeire S *et al.* Influence of immunogenicity on the long-term efficacy of infliximab in Crohn's disease. *N Engl J Med* 2003;348:601–8.
5 Anderson PJ. Tumor necrosis factor inhibitors: clinical implications of their different immunogenicity profiles. *Semin Arthritis Rheum* 2005;5 (Suppl1):19–22.
6 Elliott MJ, Maini RN, Feldmann M *et al.* Repeated therapy with monoclonal antibody to tumour necrosis factor alpha (cA2) in patients with rheumatoid arthritis. *Lancet* 1994;344:1125–7.
7 Schaeverbeke T, Truchetet ME, Kostine M, Barnetche T, Bannwarth B, Richez C. Immunogenicity of biologic agents in rheumatoid arthritis patients: lessons for clinical practice. *Rheumatology (Oxford)* 2016;55:210–20.
8 Bendtzen K. Personalized medicine: theranostics (therapeutics diagnostics) essential for rational use of tumor necrosis factor-alpha antagonists. *Discov Med* 2013;15:201–11.
9 Krieckaert CL, Nair SC, Nurmohamed MT *et al.* Personalised treatment using serum drug levels of adalimumab in patients with rheumatoid arthritis: an evaluation of costs and effects. *Ann Rheum Dis* 2015;74:361–8.
10 Lazar-Molnar E, Delgado JC. Immunogenicity assessment of tumor necrosis factor antagonists in the clinical laboratory. *Clin Chem* 2016;62:1186–98.
11 Steenholdt C, Brynskov J, Thomsen OO *et al.* Individualised therapy is more cost-effective than dose intensification in patients with Crohn's disease who lose response to anti-TNF treatment: a randomised, controlled trial. *Gut* 2014;63:919–27.
12 Thomas SS, Borazan N, Barroso N *et al.* Comparative Immunogenicity of TNF Inhibitors: impact on clinical efficacy and tolerability in the management of autoimmune diseases. A systematic review and meta-analysis. *BioDrugs* 2015;29:241–58.
13 Steenholdt C, Ainsworth MA, Tovey M *et al.* Comparison of techniques for monitoring infliximab and antibodies against infliximab in Crohn's disease. *Ther Drug Monit* 2013;35:530–8.
14 Steenholdt C, Bendtzen K, Brynskov J, Thomsen OO, Ainsworth MA. Clinical implications of measuring drug and anti-drug antibodies by different assays when optimizing infliximab treatment failure in Crohn's disease: post hoc analysis of a randomized controlled trial. *Am J Gastroenterol* 2014;109:1055–64.
15 Vande Casteele N, Buurman DJ, Sturkenboom MG *et al.* Detection of infliximab levels and anti-infliximab antibodies: a comparison of three different assays. *Aliment Pharmacol Ther* 2012;36:765–71.
16 Lallemand C, Kavrochorianou N, Steenholdt C *et al.* Reporter gene assay for the quantification of the activity and neutralizing antibody response to TNFalpha antagonists. *J Immunol Methods* 2011;373:229–39.
17 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
18 Mulleman D, Chu Miow Lin D, Ducourau E *et al.* Trough infliximab concentrations predict efficacy and sustained control of disease activity in rheumatoid arthritis. *Ther Drug Monit* 2010;32:232–6.
19 Targan SR, Hanauer SB, van Deventer SJ *et al.* A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factor alpha for Crohn's disease. Crohn's Disease cA2 Study Group. *N Engl J Med* 1997;337:1029–35.
20 Bloem K, van Leeuwen A, Verbeek G *et al.* Systematic comparison of drug-tolerant assays for anti-drug antibodies in a cohort of adalimumab-treated rheumatoid arthritis patients. *J Immunol Methods* 2015;418:29–38.
21 Jani M, Chinoy H, Warren RB *et al.* Clinical utility of random anti-tumour necrosis factor drug testing and measurement of anti-drug antibodies on long-term treatment response in rheumatoid arthritis. *Lancet* 2015;385 (Suppl 1):S48.
22 van Schouwenburg PA, Krieckaert CL, Rispens T, Aarden L, Wolbink GJ, Wouters D. Long-term measurement of anti-adalimumab using pH-shift-anti-idiotype antigen binding test shows predictive value and transient antibody formation. *Ann Rheum Dis* 2013;72:1680–6.

# II

Cytometry PART A
Journal of Quantitative Cell Science

# Titrating Complex Mass Cytometry Panels

Stein-Erik Gullaksen,[1] Lucius Bader,[2,3] Monica Hellesøy,[4] André Sulen,[1] Oda Helen Eck Fagerholt,[1] Caroline B. Engen,[1] Jørn Skavland,[1] Bjørn Tore Gjertsen,[1,4] Sonia Gavasso[5,6*]

[1]Centre of Cancer Biomarkers CCBIO, Department of Clinical Science, University of Bergen, Bergen, Norway

[2]Department of Clinical Science, University of Bergen, Bergen, Norway

[3]Bergen group of Epidemiology and Biomarkers in Rheumatic Disease (BEaBIRD), Department of Rheumatology, Haukeland University Hospital, Bergen, Norway

[4]Department of Internal Medicine, Hematology section, Helse Bergen, Bergen, Norway

[5]Department of Clinical Medicine, University of Bergen, Bergen, Norway

[6]Neuroimmunology Lab, Haukeland University Hospital, Bergen, Norway

Received 23 August 2018; Revised 26 February 2019; Accepted 28 February 2019

Grant sponsor: The Research Council of Norway, Grant number: 220759; Grant sponsor: Helse Vest health Trust, Project numbers: 912160, 912009; Grant sponsor: Norwegian Cancer Society Legacy, Grant numbers: 104712, 145268, 145269 and 163424

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Sonia Gavasso, Department of Neurology, Haukeland University Hospital, Jonas Lies Vei 65, 5021 Bergen, Norway. Email: sonia.gavasso@helse-bergen.no

Published online in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cyto.a.23751



## Abstract

We describe here a simple and efficient antibody titration approach for cell-surface markers and intracellular cell signaling targets for mass cytometry. The iterative approach builds upon a well-characterized backbone panel of antibodies and analysis using bioinformatic tools such as SPADE. Healthy peripheral blood and bone marrow cells are stained with a pre-optimized "backbone" antibody panel in addition to the progressively diluted (titrated) antibodies. Clustering based on the backbone panel enables the titration of each antibody against a rich hematopoietic background and assures that nonspecific binding and signal spillover can be quantified accurately. Using a slightly expanded backbone panel, antibodies quantifying changes in transcription factors and phosphorylated antigens are titrated on *ex vivo* stimulated cells to optimize sensitivity and evaluate baseline expression. Based on this information, complex panels of antibodies can be thoroughly optimized for use on healthy whole blood and bone marrow and are easily adaptable to the investigation of samples from for example clinical studies. © 2019 The Authors. *Cytometry Part A* published by Wiley Periodicals, Inc. on behalf of International Society for Advancement of Cytometry.

## Key terms

mass cytometry; panel design; antibody titration; whole blood; bone marrow; CyTOF; phosphoflow

MASS cytometry enables the simultaneous measurement of over 40 antigens on single cells using metal isotope conjugated antibodies, generating highly complex datasets with minimal experimental artifacts (1–3). As the number of antibodies used to investigate biologically heterogeneous cells increases, so do the demands for an efficient and thorough approach to determine optimal antibody titers. In addition to undesirable signal arising from nonspecific antibody binding, which is an issue in all types of antibody-based assays, three sources of signal spillover exist in mass cytometry (4,5): signal overlap of highly abundant metal isotopes into adjacent mass channels (±1 Da), isotope oxidation (+16 Da), and isotopic impurities in the metal isotopes. Although technical approaches to deal with similar experimental artifacts have been well established for conventional flow cytometry (6), mass cytometry has unique requirements (7).

The predictable patterns of signal spillover in mass cytometry are not routinely compensated, as is commonplace in conventional flow cytometry. Although such compensation tools have been developed (8), signal spillover can be significantly reduced by lowering the signal intensities (linearly dependent) and/or by carefully designing antibody panels (9). The former may not allow for sufficient distance between the biologically positive and dim/negative populations, and the latter may introduce unwanted/unnecessary noise in the data. In contrast to conventional flow cytometry the range of "brightness" observed across the mass range of purified metal isotope tags is fairly equal (1,10). Thus, the choice of isotope may not always provide

the staining characteristics needed to capture the biological diversity within the mass cytometers dynamic range, without accepting signal spillover to some degree. Antibody binding to secondary and low affinity epitopes must also be evaluated. This might be a challenging process as the combinatorial possibilities of marker expression quickly exceeds our understanding of the human immune system with increasing numbers of markers. Lastly, we emphasize that determination of optimized antibody titer is application-specific and is not necessarily transferrable between different biological samples, processing protocols, laboratories, or antibody lots. In addition, we have also observed a variation in the stability of metal conjugated antibodies, potentially changing the optimal titer over time. Taken together, the construction of large antibody panels for mass cytometry is an extremely time consuming, laborious, and demanding undertaking, necessitating an efficient and straightforward approach.

## Materials and Methods

### Subjects and Samples

Peripheral blood (PB) and Bone marrow (BM) samples were obtained from healthy individuals who provided written informed consent (local ethical committee approval 2012/2247). PB and BM were collected in the presence of heparin. The leukocytes were fixed and erythrocytes lysed using Lyse/Fix buffer (BD Biosciences) within 1 h, and samples were stored at −80°C in physiological saline.

### *Ex Vivo* Stimulation of Peripheral Blood and Bone Marrow

Freshly collected PB and BM from one healthy donor were stimulated *ex vivo* with IFN-α (100 ng/ml, 15 min), GM-CSF (100 ng/ml, 15 min), LPS (10 μg/ml, 15 min), or left untreated. PB and BM cells were fixed, and erythrocytes lysed using the BD Lyse/Fix reagent as above.

### Barcoding and Antibody Staining

Fixed leukocytes from PB and BM were barcoded (3) using the Fluidigm 20-plex metal barcoding kit according to manufacturer's protocol. All antibodies used in this study were either purchased pre-conjugated from Fluidigm or were conjugated in-house using the X8 MaxPar conjugation kits according to manufacturer's protocol (Online Tables 1–3). See the online materials for detailed protocols. Briefly, aliquots of $1.5 \times 10^6$ cells were first pretreated with heparin (100 IU/ml, 20 min) (11) and then stained with mastermixes of backbone antibody panel mixed with twofold serially diluted panel of antibodies to be titrated in a total staining volume of 50 μl (30 min, room temperature). The dilution of most antibodies started at the concentration recommended by the manufacturer (1 μl antibody per 100 μl cell suspension containing $3 \times 10^6$ cells). However, for some antibodies, a pre-dilution was necessary before a twofold titration was possible. For instance, the $^{163}$Dy-CD56 was diluted by a factor of 10× (Online Fig. 8) before the twofold dilution shown in Supplemental Figure 2. Cells to be stained with intracellular signal transduction antibodies were permeabilized for 10 min on ice with methanol (−20°C, 100%), treated with heparin (100 IU/ml, 20 min) and subsequently stained with progressively titrated (five twofold dilutions) intracellular antibodies (30 min, room temperature). To enable the identification of cells, the DNA was labeled with iridium-191/193 by incubation in 0.1 nM Ir-nucleic acid intercalator (Fluidigm) diluted in MaxPar PBS containing 4% PFA (Alfa Aesar, 16% PFA, methanol-free) overnight at 4°C. Cells that were not permeabilized with methanol (cell surface only) were labeled with iridium-191/193 by incubating with Ir-nucleic acid intercalator (0.1 nM) diluted in MaxPar Fix/perm buffer (Fluidigm) overnight at 4°C. Immediately before sample acquisition, cells were washed in MaxPar cell staining buffer and MaxPar water (both from Fluidigm) and left pelleted until analysis on the Helios mass cytometer (Fluidigm). The cells were resuspended in MaxPar water supplemented with a 1:10 dilution of the EQ Four Element calibration beads (Fluidigm). The acquisition rate was kept below 400 cells per second to limit the number of acquired cell doublets.

### Data Analysis

Machine drift in the data was normalized using the Fluidigm bead normalizer. Cell debris and doublets were manually removed by gating on event length and DNA (Ir-191/193). The Fluidigm barcode de-convolution tool was used for de-barcoding samples. The histogram overlay illustrations were made, and SPADE (12) analysis was performed, in cytobank.org. Sample concatenation and gating was performed in FlowJo (FLOWJO, LLC). For gated populations, the 75th percentile of the dual count in each mass channel was exported for statistics. The heat maps were made using Morpheus (https://software.broadinstitute.org/morpheus/).

## Results and Discussion

A graphical illustration of our approach is presented in Figure 1, and a more detailed description is given in the Online Materials. All reagents used in this work can be found in Online Tables 1–3. A backbone panel (Online Table 4) of carefully selected antibodies was established as basis to evaluate the titration of additional antibodies in titration step 1. Titration data for the backbone panel is shown in Figure 1a and Online Figure 1a. The optimal titer for each antibody (red gate) was approximated by contrasting the ability to securely discern positive from negative cells against signal spillover into other mass channels. For example, low-level spillover of 2–3 dual counts of $^{145}$Nd-CD4 signal into $^{146}$Nd-CD8 can be accepted, as long as co-expression of CD8 is not of biologic interest. In titration step 2, after optimization of the backbone panel, sample aliquots of metal barcoded (3) and paired PB and BM from two healthy donors were stained with the backbone panel and serially diluted mastermixes of the three "titration panels" (Online Table 6) containing additional cell surface antibodies (Fig. 1b). In these titration panels, all channels theoretically receiving spillover from the included markers were kept empty. For example, $^{144}$Nd-CD38 and $^{148}$Nd-CD16 were placed in different titration panels. We used the SPADE

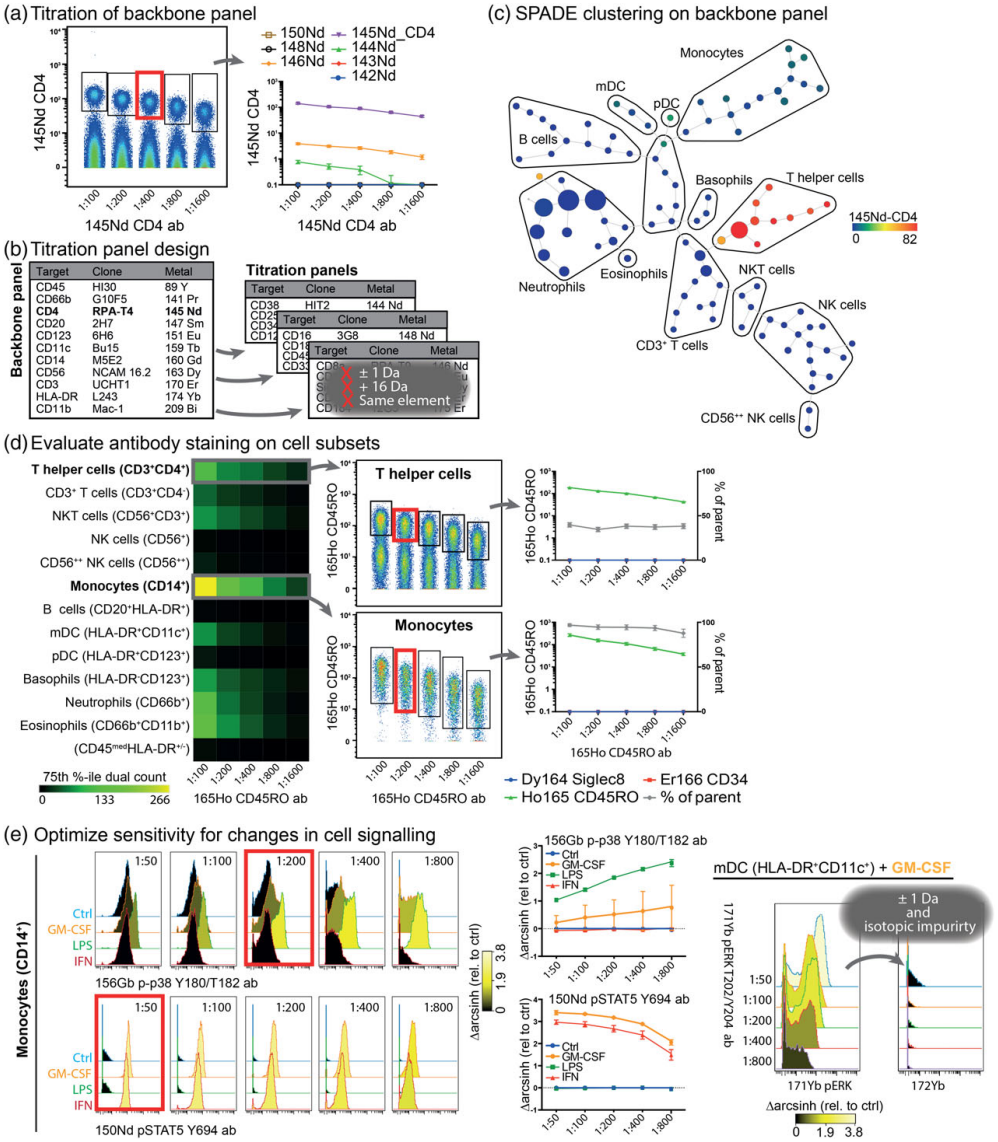



**Figure 1.** (a) A backbone panel was titrated using PB and BM from two healthy donors. The .fcs files were concatenated to visualize immune staining as a function of antibody concentration and to enable efficient gating of positive cells. The expression of the titrated antigen and all spillovers (±1 Da, +16 Da and channels detecting isotopic impurities) was calculated for the gated cells plotted (75th %-ile dual counts). Optimal titer (red gate) was chosen by balancing the ability to discern positive from negative cells with the amount of signal overlap generated in other mass channels. (b) Additional cell surface antibodies to be titrated were subdivided into titration panels. Here, all channels receiving spillover were unused in each panel. Cell sample aliquots were stained with the titrated backbone panel and serially diluted mastermixes of the titration panels. (c) A single SPADE clustering was performed to efficiently identify cell subsets in the entire data set. The clustering was based solely on the backbone panel, and cell subsets manually identified. (d) The signal from the titrated antibodies were measured in each of the cell subsets and plotted as a heat map. The data was in selected cell subsets concatenated, and the expression of the titrated antigen and all spillovers calculated for the gated population, as above. The exact staining pattern on a relevant cell subset (i.e., CD45RO expression on T helper cells) could now also be evaluated in addition to signal spillover (i.e., CD45RO expression on monocytes) and panel design. The red gate indicates the chosen antibody titer. The relative abundance of positive cells in the parent cell subset as a function of antibody concentration was also calculated. (e) PB and BM from one healthy donor were stimulated *ex vivo* with GM-CSF (100 ng/ml, 15 min), IFN-α (100 ng/ml, 15 min) or LPS (10 μg/ml, 15 min) or left untreated. The antibodies to be titrated were split into two titration panels, as above. Cells were stained with backbone panel and serially diluted titration panels, and cell subsets identified using SPADE. The phosphorylation level (75th %-ile) was measured in each population, for all intracellular antibodies and all channels theoretically receiving spillover. Drug-induced changes in phosphorylation levels were calculated (Δarcsinh relative to ctrl) and plotted as a function of antibody dilution. Lastly, the signal spillover generated by induction of signaling into the empty mass channel was evaluated. Red boxes indicate optimal dilutions of antibodies. Color scales indicate Δarcsinh relative to control.

**Table 1.** Antibody panel. (See online Tables 1–4, 6, and 8 in the online materials for more details)

| SPECIFICITY | CLONE | ISOTOPE | PURPOSE |
|---|---|---|---|
| CD45 | HI30 | 89 Y | Pan leukocytes |
| CD66b | G10F5 | 141 Pr | Neutrophils |
| Cleaved caspase 3 | D3E9 | 142 Nd | Apoptosis |
| CD38 | HIT2 | 144 Nd | Activation |
| CD4 | RPA-T4 | 145 Nd | T helper cells |
| CD8a | RPA-T8 | 146 Nd | Cytotoxic T cells |
| CD20 | 2H7 | 147 Sm | B cells |
| CD16 | 3G8 | 148 Nd | Neutrophils and subsets of NK and monocytes |
| CD25 | 2A3 | 149 Sm | Basophils, Tregs, and activated T helper cells |
| pSTAT5 Y694 | 47 | 150 Nd | Signal transduction |
| CD123 | 6H6 | 151 Eu | Basophils, mDC, and pDC |
| pSTAT1 Y701 | 58D6 | 153 Eu | Signal transduction |
| p-p38 T180/Y182 | D3F9 | 156 Gd | Signal transduction |
| pSTAT3 Y705 | 4/P-STAT3 | 158 Gd | Signal transduction |
| CD11c | Bu15 | 159 Tb | Monocytes and mDC |
| CD14 | M5E2 | 160 Gd | Monocytes |
| CD181 (IL-8RA) | B1 | 161 Dy | Neutrophils |
| FoxP3 | PCH101 | 162 Dy | Tregs |
| CD56 | NCAM 16.2 | 163 Dy | NK cells |
| CD45RO | UCHL1 | 165 Ho | Naïve/memory T cells |
| CD34 | 581 | 166 Er | Hematopoietic stem/progenitor cell |
| CD1c (BDCA-1) | L161 | 167 Er | Subsets of mDC and B cells |
| CD335 (NKp46) | 9E2 | 169 Tm | NK cells |
| CD3 | UCHT1 | 170 Er | T cells |
| pERK 1/2 T202/Y204 | D1314.E4 | 171 Yb | Signal transduction |
| HLA-DR | L243 | 174 Yb | Activation, DCs, monocytes, and B cells |
| CD184 (CXCR4) | 12G5 | 175 Lu | Basophils |
| pCREB S133 | 87G3 | 176 Yb | Signal transduction |
| CD11b | Mac-1 | 209 Bi | Granulocytes, monocytes NK cells, and DCs |

mDC; myeloid dendritic cell; pDC, plasmacytoid dendritic cell; NK, natural killer; Tregs, regulatory T cells.

clustering algorithm (12) in Cytobank.org to identify common cell subsets across all files in the experiment based on backbone antigen expression (Fig. 1c and Online Fig. 2). This clustering provided a rich immune-phenotypic background on which the titration of the additional antibodies could be evaluated (Fig. 1d and Online Fig. 1b–d). In addition to evaluating signal spillover using the cell population with the highest expression (e.g., CD45RO on monocytes) as in step 1, this also allowed for the exact evaluation of the staining pattern in a biologically relevant cell subset (e.g., T helper cells). In this way, staining characteristics can be seen across a wider hematological background, altogether further refining the approximation of optimal titers (indicated in red boxes). In titration step 3, we evaluated antibodies specific for intracellular cell signaling and transcription factors. A metal barcoded pool of *ex vivo* stimulated PB and BM (IFN-α [100 ng/ml, 15 min], GM-CSF [100 ng/ml, 15 min], LPS [10 μg/ml, 15 min]) was stained with a combined backbone panel based on titration steps 1 and 2, and progressively diluted titration panels as described above (Online Table 8). After SPADE clustering on surface antigen expression as above (Online Figs. 3 and 4), we calculated the stimulation-induced change in cell signaling (Δarcsinh relative to control) for all cell subsets (Fig. 1e and Online Fig. 5). Of note, in our experiment the Δarcsinh after both GM-CSF and LPS stimulation increased for p-p38 Y180/T182 in monocytes ($CD14^+$) with increasing dilution of the antibody. Likely, surplus antibody created an increased background, thus masking a drug-induced regulation in signal transduction after stimulation. This emphasizes the importance of selecting optimal antibody titers using appropriate biological controls. Furthermore, we assessed the signal spillover as a function of drug-induced alterations in cell signaling. For example, after GM-CSF stimulation, we could measure spillover signal into the empty $^{172}Yb$ channel induced by high phosphorylation levels of pERK1/2 Y202/T204 ($^{171}Yb$) in the myeloid dendritic cell population (mDCs, $CD11c^+HLA\text{-}DR^+$). This spillover decreased as a function of antibody titration (Fig. 1e, right panel). The final choice of antibody titers was done by minimizing signal spillover and optimal resolution between positive/stimulated and negative/baseline. We validated our approach by testing the titrated panel (Table 1 and Online Table 1 and 2) on three additional healthy donor PB samples. The staining patterns of both cell surface markers and intracellular signal transduction targets in these additional samples reproduced the antibody

titration results, highlighting the usefulness of our approach (See online materials and Online Figs. 6 and 7).

In summary, we outline here a conceptual framework where we highlight the usefulness of performing iterative antibody titration on cells stained with a backbone panel. We found SPADE to be an excellent tool for automated cell clustering based on the backbone panel. SPADE enabled clustering of cells in a dataset consisting of more than 6 million cells into a single SPADE tree. Using bioinformatic tools, this approach is efficient and straightforward and provides a deeper characterization of each antibody's performance, which is necessary for the demanding task of panel design for mass cytometry assays. Although we have demonstrated the titration of antibodies on healthy PB and BM in this work, this approach can easily be adapted to other sample types for mass cytometry.

## Acknowledgments

This study was supported by The Research Council of Norway (Petromaks program grant #220759), Helse Vest Health Trust (project no. 912160, 912009), and the Norwegian Cancer Society Legacy (Grant no. 104712, 145268, 145269 and 163424) with Solveig & Ole Lunds Legacy. The Helios mass cytometry instrument was a generous gift from the Trond Mohn Foundation.

## Literature Cited

1. Spitzer MH, Nolan GP. Mass cytometry: Single cells, many features. Cell 2016;165(4):780–791.
2. Catena R, Ozcan A, Zivanovic N, Bodenmiller B. Enhanced multiplexing in mass cytometry using osmium and ruthenium tetroxide species. Cytometry A 2016;89(5):491–497.
3. Zunder ER, Finck R, Behbehani GK, Amir el AD, Krishnaswamy S, Gonzalez VD, et al. Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. Nat Protoc 2015;10(2):316–333.
4. Ornatsky OI, Kinach R, Bandura DR, Lou X, Tanner SD, Baranov VI, Nitz M, Winnik MA. Development of analytical methods for multiplex bio-assay with inductively coupled plasma mass spectrometry. J Anal At Spectrom 2008;23(4):463–469.
5. Bendall SC, Nolan GP, Roederer M, Chattopadhyay PK. A deep profiler's guide to cytometry. Trends Immunol 2012;33(7):323–332.
6. Mahnke YD, Roederer M. Optimizing a multicolor immunophenotyping assay. Clin Lab Med 2007;27(3):469–485.
7. Takahashi C, Au-Yeung A, Fuh F, Ramirez-Montagut T, Bolen C, Mathews W, O'Gorman WE. Mass cytometry panel optimization through the designed distribution of signal interference. Cytometry A 2017;91(1):39–47.
8. Chevrier S, Crowell HL, Zanotelli VRT, Engler S, Robinson MD, Bodenmiller B. Compensation of signal spillover in suspension and imaging mass cytometry. Cell Syst 2018;6(5):612–620.
9. Leipold MD, Newell EW, Maecker HT. Multiparameter phenotyping of human PBMCs using mass cytometry. Methods Mol Biol 2015;1(343):81–95.
10. Tricot S, Meyrand M, Sammicheli C, Elhmouzi-Younes J, Corneau A, Bertholet S, Malissen M, le Grand R, Nuti S, Luche H, et al. Evaluating the efficiency of isotope transmission for improved panel design and a comparison of the detection sensitivities of mass cytometer instruments. Cytometry A 2015;87(4):357–368.
11. Rahman AH, Tordesillas L, Berin MC. Heparin reduces nonspecific eosinophil staining artifacts in mass cytometry experiments. Cytometry Part A 2016;89(6):601–607.
12. Qiu P, Simonds EF, Bendall SC, Gibbs KD Jr, Bruggner RV, Linderman MD, et al. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nat Biotechnol 2011;29(10):886–891.

# III

frontiers in Immunology

ORIGINAL RESEARCH
published: 05 July 2019
doi: 10.3389/fimmu.2019.01488

# Candidate Markers for Stratification and Classification in Rheumatoid Arthritis

*Lucius Bader*[1,2]*, *Stein-Erik Gullaksen*[3,4], *Nello Blaser*[5], *Morten Brun*[5], *Gerd Haga Bringeland*[6,7], *André Sulen*[1,6], *Clara Gram Gjesdal*[1,2], *Christian Vedeler*[6,7] and *Sonia Gavasso*[1,6,7]*

[1] Bergen Group of Epidemiology and Biomarkers in Rheumatic Disease, Department of Rheumatology, Haukeland University Hospital, Bergen, Norway, [2] Department of Clinical Science, University of Bergen, Bergen, Norway, [3] Center of Cancer Biomarkers, University of Bergen, Bergen, Norway, [4] Department of Internal Medicine, Hematology Section, Haukeland University Hospital, Bergen, Norway, [5] Department of Mathematics, University of Bergen, Bergen, Norway, [6] Department of Neurology, Haukeland University Hospital, Bergen, Norway, [7] Department of Clinical Medicine, University of Bergen, Bergen, Norway

OPEN ACCESS

**Edited by:**
Helen Marie McGuire,
University of Sydney, Australia

**Reviewed by:**
João Eurico Fonseca,
University of Lisbon, Portugal
Sinisa Savic,
University of Leeds, United Kingdom
Amy Elizabeth Anderson,
Newcastle University, United Kingdom

***Correspondence:**
Lucius Bader
lucius.bader@helse-bergen.no
Sonia Gavasso
sonia.gavasso@helse-bergen.no

**Specialty section:**
This article was submitted to Autoimmune and Autoinflammatory Disorders, a section of the journal Frontiers in Immunology

**Received:** 15 March 2019
**Accepted:** 14 June 2019
**Published:** 05 July 2019

**Citation:**
Bader L, Gullaksen S-E, Blaser N, Brun M, Bringeland GH, Sulen A, Gjesdal CG, Vedeler C and Gavasso S (2019) Candidate Markers for Stratification and Classification in Rheumatoid Arthritis. Front. Immunol. 10:1488. doi: 10.3389/fimmu.2019.01488

Rheumatoid arthritis (RA) is a chronic autoimmune, inflammatory disease, characterized by synovitis in small- and medium-sized joints and, if not treated early and efficiently, joint damage, and destruction. RA is a heterogeneous disease with a plethora of treatment options. The pro-inflammatory cytokine tumor necrosis factor (TNF) plays a central role in the pathogenesis of RA, and TNF inhibitors effectively repress inflammatory activity in RA. Currently, treatment decisions are primarily based on empirics and economic considerations. However, the considerable interpatient variability in response to treatment is a challenge. Markers for a more exact patient classification and stratification are lacking. The objective of this study was to identify markers in immune cell populations that distinguish RA patients from healthy donors with an emphasis on TNF signaling. We employed mass cytometry (CyTOF) with a panel of 13 phenotyping and 10 functional markers to explore signaling in unstimulated and TNF-stimulated peripheral blood mononuclear cells from 20 newly diagnosed, untreated RA patients and 20 healthy donors. The resulting high-dimensional data were analyzed in three independent analysis pipelines, characterized by differences in both data clean-up, identification of cell subsets/clustering and statistical approaches. All three analysis pipelines identified p-p38, IkBa, p-cJun, p-NFkB, and CD86 in cells of both the innate arm (myeloid dendritic cells and classical monocytes) and the adaptive arm (memory $CD4^+$ T cells) of the immune system as markers for differentiation between RA patients and healthy donors. Inclusion of the markers p-Akt and CD120b resulted in the correct classification of 18 of 20 RA patients and 17 of 20 healthy donors in regression modeling based on a combined model of basal and TNF-induced signal. Expression patterns in a set of functional markers and specific immune cell subsets were distinct in RA patients compared to healthy individuals. These signatures may support studies of disease pathogenesis, provide candidate markers for response, and non-response to TNF inhibitor treatment, and aid the identification of future therapeutic targets.

**Keywords: rheumatoid arthritis, patient stratification, tumor necrosis factor, tumor necrosis factor inhibitors, mass cytometry**

## INTRODUCTION

Rheumatoid arthritis (RA) is a chronic autoimmune disease characterized by synovial inflammation that, if not treated early and efficiently, causes joint damage. The pro-inflammatory cytokine tumor necrosis factor (TNF) plays a central role in the pathogenesis of RA and is the target of treatment with TNF inhibitors. TNF inhibitors are generally effective and well-tolerated (1, 2); however, up to one-third of patients are primary non-responders, and responses in up to one-third of initial responders abate over time (3, 4).

Currently, only a few markers for diagnostic and stratification purposes are used in daily clinical practice in patients with RA. Anti-citrullinated peptide antibodies are a highly disease-specific biomarker with an impact mostly on diagnosis and classification (5). TNF inhibitor drug levels and anti-drug antibodies are indicative of treatment responses; however, these markers are not standardized for clinical application (3, 6, 7). Several candidate biomarkers for prediction of treatment responses have been suggested based on gene, cytokine, and immune cell profiles, but none have added significant value to patient stratification in a clinical setting (8). Previous studies have indicated the potential of single-cell profiling by flow or mass cytometry in patient stratification in RA and in other autoimmune conditions (9, 10). Distinct signaling patterns have been found in RA patients before and during treatment with TNF inhibitors in exploratory and proof-of-principle studies (11, 12).

We hypothesize that signaling patterns in RA are distinct from those of healthy donors. The unbiased identification of RA-specific signaling patterns in immune cell subsets before treatment may improve diagnosis, therapeutic stratification, and monitoring, and may also facilitate studies of disease pathogenesis and the development of drugs that target dysfunctional pathways with high precision.

In this study, we used mass cytometry to explore signaling responses to TNF in single immune cells of RA patients and healthy donors. In mass cytometry metal-tagged antibodies serve as markers with a read-out in a mass spectrometry time-of-flight chamber (13). Using mass cytometry, up to 50 markers can be simultaneously analyzed with single cell resolution with relatively little signal overlap and very low background noise (14, 15). Here we used a panel of 13 phenotyping and 10 functional markers for an in-depth characterization of peripheral blood mononuclear cells (PBMCs) from patients and controls with and without stimulation with TNF. Based on results from three different analysis pipelines, we suggest a smaller set of phenotyping and functional markers, which strongly correlate with disease status for future use in e.g., flow cytometry.

## MATERIALS AND METHODS

In-depth information on material, methods and results is provided in the **Supplementary Material** in the same order and with the same headings/sub-headings as in the main article.

### Healthy Donors and RA Patients

Twenty healthy donors (HD, 4 male, 16 female, ages 39–67) and 20 RA patients (4 male, 16 female, ages 31–76) were included in this study (**Table 1**). All RA patients were included at the time of diagnosis and fulfilled the ACR/EULAR 2010 criteria for RA. None of the patients had received synthetic or biologic disease-modifying anti-rheumatic drugs, but five had been prescribed low to moderate dosages of prednisolone by their general practitioners prior to the first consultation with a rheumatologist. Despite ongoing prednisolone-treatment at inclusion (range 2.5–15 mg), these patients had high disease activity with a mean disease activity score (DAS28) of 6.1 (range 5.4–7).

All donors and patients gave written informed consent for inclusion into the Norwegian Arthritis Registry (NorArtritt) and the Research Biobank for Rheumatic Diseases in Western Norway (approval REK 2012/1689). Utilization of registry data and biobank material for this study was approved by the Regional Ethics Committee (approval REK 2014/317).

### Peripheral Blood Mononuclear Cells (PBMCs)

PBMCs were chosen due to the possibility of culturing and application of standardized and simultaneous conditions (such as e.g., cytokine stimulation) after cryo-preservation.

PBMCs were harvested by density gradient centrifugation (BD Vacutainer® CPT™ Mononuclear Cell Preparation Tube—Sodium Citrate), processed for cryo-preservation within 4 h and stored in liquid nitrogen in 50% hematopoietic cell medium (X-VIVO™, Lonza), 42.5% freezing medium (ProFreeze™, Lonza), and 7.5% dimethyl sulfoxide (Sigma-Aldrich).

**TABLE 1** | Patient and healthy donor characteristics.

| | |
|---|---|
| **Healthy donors (HD)** | |
| Female/male | 16/4 |
| Median age (range) | 49 (34–67) years |
| **RA patients (RA)** | |
| Female/male | 16/4 |
| Median age (range) | 63.5 (31–76) |
| **Disease characteristics** | |
| RF+ | 14 |
| ACPA+ | 11 |
| RF+ ACPA+ | 9 |
| RF- ACPA- | 4 |
| Mean DAS28 (range) | 5.37 (3–7.6) |
| Mean DAS28–CRP (range) | 4.86 (2.5–7.2) |
| Mean CRP (range) | 25.3 (1–156) |
| Mean ESR (range) | 36.6 (6–104) |
| **Medication** | |
| Prednisolone | 5 of 20 patients |
| Prednisolone daily dose | 2.5-15 mg (2.5, 2.5, 12.5, 12.5, 15 mg) |

*RF, rheumatoid factor; ACPA, anti-citrullinated peptide antibodies; DAS28, disease activity score with 28 joint count; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate.*

## Antibody Panel

All antibodies used in this study (**Table 2**) were titrated on PBMCs from one healthy donor. Titrations were performed on unstimulated PBMCs and cells stimulated with TNF and phorbol myristate acetate (PMA) for optimization of pathway activation markers. The antibodies against CD120a, CD120b, and p-cJun were conjugated to metals in our laboratory (conjugation kits and protocols by Fluidigm), all other antibodies were pre-conjugated (Fluidigm).

Thirteen markers were applied to define common PBMC subsets; these were used in both automated clustering and manual gating. Functional markers for TNF signaling were the cleaved caspase 3 as a marker for apoptosis signaling; p-p38 [T180/Y182] and p-Erk1/2 [T202/Y204] as markers for the MAPK-pathway activation; IkBa and p-NFkB p65 [S529] for the NFkB canonical pathway; p-Akt [S473] for the PI3K-Akt pathway; and p-cJun [S73] for the SAPK/JNK signaling pathway. CD86 was added as a marker of T cell regulation and analyzed as functional marker, although signaling through this pathway is not directly related to TNF.

Treatment with PMA resulted in significant increases in Erk1/2 phosphorylation, whereas TNF treatment did not have significant effects on Erk1/2 phosphorylation. This marker was therefore omitted from experiments after panel titration.

**TABLE 2 |** Antibody panel with epitopes, antibody clones, conjugated metals, and target cell populations or signaling pathways.

| Epitope | Clone | Metal | Target/Function | Abbrev. |
|---|---|---|---|---|
| **PHENOTYPING** | | | | |
| CD20 | 2H7 | $^{147}$Sm | B lymphocytes | Bc |
| CD3 | UCHT1 | $^{170}$Er | T lymphocytes | |
| CD4 | RPA-T4 | $^{145}$Nd | CD4$^+$ T lymphocytes | CD4 Tc |
| CD8a | RPA-T8 | $^{146}$Nd | CD8$^+$ T lymphocytes | CD8 Tc |
| CD45RA | HI100 | $^{169}$Tm | Naïve/effector vs. memory | Naïve, mem |
| CD56 | NCAM16.2 | $^{176}$Yb | Natural killer cells | NKc |
| CD16 | 3G8 | $^{148}$Nd | NK T cells | NK Tc |
| CD14 | M5E2 | $^{160}$Gd | Classical monocytes | cM |
| CD61 | VI-PL2 | $^{209}$Bi | Monocytes | cM |
| CD11c | Bu15 | $^{159}$Tb | Myeloid dendritic cells | mDc |
| CD123 (IL-3R) | 6H6 | $^{151}$Eu | Plasmacytoid dendritic cells | pDc |
| HLA-DR | L243 | $^{174}$Yb | MHCII, antigen presentation | |
| CD45 | HI30 | $^{89}$Y | Leukocyte Common Antigen | |
| **FUNCTIONAL** | | | | |
| Cleaved Caspase 3 | D3E9 | $^{142}$Nd | Apoptotic signaling | Caspase3 |
| p-p38 [T180/Y182] | D3F9 | $^{156}$Gd | MAPK pathway | p-p38 |
| p-Erk1/2 [T202/Y204]* | D13.14.4E | $^{171}$Yb | MAPK pathway | p-Erk |
| p-Akt [S473] | D9E | $^{152}$Sm | PI3K-Akt pathway | p-Akt |
| p-cJun [S73]** | D47G9 | $^{167}$Er | SAPK/JNK signaling | p-cJun |
| p-NFkB p65 [S529] | K10-895.12.50 | $^{166}$Er | NFkB canonical pathway | p-NFkB |
| IkBa | L35A5 | $^{164}$Dy | with IkBa degradation | IkBa |
| CD120a** | MABTNFR1-B1 | $^{155}$Gd | TNF receptor 1 | TNFR1 |
| CD120b** | hTNR-M1 | $^{165}$Ho | TNF receptor 2 | TNFR2 |
| CD86 | IT2.2 | $^{150}$Nd | Regulation of T cell activity | CD86 |

**p-Erk1/2 was omitted from the panel after TNF titration experiments, since TNF stimulation did not alter p-Erk1/2 expression.*

***Metal-conjugation carried out at our laboratory (all other antibodies were purchased pre-conjugated).*

## Experimental Workflow

Cryopreserved PBMCs were thawed, rapidly transferred to warm X-VIVO™ containing a nuclease (Benzonase® Nuclease, Merck Millipore, 25 U/mL), followed by centrifugation and resting in X-VIVO™ for 4 h at 37°C, 5% $CO_2$. The resting time was optimized in set-up experiments (data not shown). Viability staining was performed according to the manufacturer's instructions with Cell-ID™ cisplatin (Fluidigm). PBMCs from each individual were split into two aliquots; one was not stimulated, and the other was stimulated with 50 ng/mL TNF for 12 min. Stimulation time and dose had been defined after a series of TNF time and dose titrations. Cells in both samples were fixed in proteomic stabilizer (Smart Tubes Inc.) for 10 min and stored at −80°C until barcoding and staining. All cells were barcoded simultaneously with 20-plex Cell-ID™ barcoding kits (Fluidigm) as recommended by the manufacturer. After pooling, surface staining, methanol-permeabilization, and intracellular staining were carried out. PBMCs were then stained with MaxPar DNA intercalator overnight (Fluidigm) and analyzed the following day on a Helios mass cytometer (Fluidigm) after addition of normalization beads (Fluidigm). Raw FCS-files were bead-normalized, concatenated, and debarcoded with software tools from Fluidigm before subsequent analysis.

## Data Analysis Workflow

Three independent analysis pipelines were performed to test and share different approaches as well as to validate our in-house NM2B algorithm (pipeline 1).

### Pipeline 1: NM2B Algorithm

This algorithm consisted of three main steps: preprocessing (A), finding cell types (B) and classification (C).

A. Preprocessing: We fitted a mixture of two Gaussian distributions to mean-variance scaled "Event_length," "Center," "Offset," "Width," "Residual," "191Ir_DNA1," "193Ir_DNA2" markers for cleanup and discarded data belonging to the smaller cluster as doublets and debris (16, 17).
B. Finding cell types: We used the following phenotyping markers to detect cell types: "147Sm_CD20," "170Er_CD3," "145Nd_CD4," "146Nd_CD8a," "169Tm_CD45RA," "176Yb_CD56," "148Nd_CD16," "160Gd_CD14," "209Bi_CD61," "159Tb_CD11c," "151Eu_CD123," and

"174Yb_HLA-DR." We performed farthest point sampling to find 49 clusters. Farthest point sampling is an approximation to k-means clustering, which can be calculated for large datasets. Clusters of size less than 1/100,000 of the total data size were discarded. We then employed complete linkage meta-clustering of the farthest points with 15 meta-clusters and discarded all meta-clusters of less than 0.5% of the total data size. The final 12 meta-clusters contained a total of 18,374,011 cell events (**Supplementary Figure 3**).

C. Classification: We used the following functional markers as features for classification of individuals as patients or controls: "142Nd_Caspase3," "156Gd_p-p38," "152Sm_p-Akt," "167Er_p-cJun," "166Er_p-NFkB," "164Dy_IkBa," "155Gd_CD120a," "165Ho_CD120b," "150Nd_CD86." For each meta-cluster we calculated the median and 90% quantile of each of the functional markers for all basal cells. In addition, we calculated the arcsinh ratios of the expression of functional markers in stimulated and basal cells. We tested three models, based on either only basal variables (basal), or only arcsinh ratios between stimulated and basal variables (ratio) or both basal and arcsinh ratios (combined). We fitted a logistic lasso regression model, that is a logistic regression model with automatic variable selection, using double leave-one-out cross validation. For details of how double leave-one-out cross-validation was performed we refer to the **Supplementary Material**. We report cross-validation accuracy, area under the ROC curve (AUC), and all non-zero coefficients.

### Pipeline 2: CITRUS Algorithm

Normalized, concatenated and debarcoded files were imported in Cytobank for downstream analysis (18). Data were cleaned for doublets, debris and dead cells by biaxial gating and analyzed with the cluster identification, characterization, and regression tool CITRUS in cytobank.org, applying the predictive regression model Nearest Shrunken Centroid/PAMR (19). CITRUS was run on the same data set, but with independent downsampling, with 3 repetitions.

### Pipeline 3: Manual Analysis

Normalized, concatenated, and debarcoded files were imported in Cytobank. Data were cleaned for doublets, debris, and dead cells by biaxial gating. viSNE analysis based on t-distributed stochastic neighbor embedding was performed for each donor and patient after downsampling to 50,000 cell events per individual and condition (20), and cell subsets were gated on individual viSNE plots (**Supplementary Figure 6**). Expression of functional markers was compared in all cell subsets, both unstimulated and TNF-stimulated, by applying non-parametric Mann-Whitney U tests using GraphPad prism version 7.0c for Mac OS X. A correction for multiple comparisons was not conducted due to the explorative character of this study.

## RESULTS

### Pipeline 1: NM2B Algorithm

Single-cell data from all 40 individuals were clustered and meta-clustered, and different numbers of clusters and meta-clusters were tested. The model used provided the best translation of meta-clusters into common immune cell subsets. The numbers of clusters and meta-clusters influenced cross-validation accuracy for the classification of RA patients and healthy donors, but when higher numbers of meta-clusters were used, the results were more difficult to interpret with regards to common cell subsets (data not shown).

Results presented here are based on 49 clusters and 12 meta-clusters; the latter include one B cell meta-cluster (4.3%), four of T cells (75.3%), two of natural killer cells (5%), one of classical monocytes (5.7%), three of myeloid dendritic cells (8.8%) and one of plasmacytoid dendritic cells (1%). Phenotyping markers are differentially expressed in the meta-clusters (**Figure 1A**); differences in expression of phenotyping markers in healthy donors vs. RA patients were not significant (**Supplementary Figure 10**).

A regression model based on both basal expression of functional markers and arcsinh ratios ("combined model") provided the best predictive TNF signaling patterns for healthy donors and RA patients. In this model seven functional markers (IkBa, CD120b, CD86, p-cJun, p-NFkB, p-p38, and p-Akt) in five cell subsets (memory $CD4^+$ T cells, $CD11c+HLA\text{-}DR+CD14^{low}CD61^{low}$ myeloid dendritic cells, naïve $CD4^+CD45RA^+CD11c^{low}$ T cells, classical monocytes, and $CD11c^{high}HLA\text{-}DR^{high}CD61^{low}$ myeloid dendritic cells) were identified as predictive markers (**Figure 1B**). Applying these markers, the combined model correctly classified 18 of 20 RA patients and 17 of 20 healthy donors (**Figure 1C**). The two patients who were not classified correctly were both females older than 67 years with high disease activity (DAS28 5.3/6.4). One was seronegative and one was being treated with prednisolone at 12.5 mg per day. Principle component analysis (PCA) of features identified by the Lasso-regression showed a good separation of HD vs. RA in the combined (**Figure 1D**) and basal model, but to a lesser degree in the ratio model (**Supplementary Figures 7–9**).

### Pipeline 2: CITRUS Algorithm

We performed four repetitive CITRUS analyses of basal expression of functional markers. p-p38, IkBa, p-cJun, p-NFkB, and CD86 were identified as predictive markers by CITRUS, with memory $CD4^+$ T cells being the most relevant cell subset for both p-p38, IkBa, and p-cJun, while clusters within myeloid dendritic cell subets (mDc) and classical monocytes (cM) were the most relevant for p-NFkB and CD86 (**Supplementary Figure 11** and **Supplementary Table 10**). There was not always a clear distinction between myeloid dendritic cells and classical monocytes in hierarchical clustering in CITRUS, and both these cell subsets were relevant for the markers p-NFkB and CD86 (**Figure 1E** and **Supplementary Table 10**).

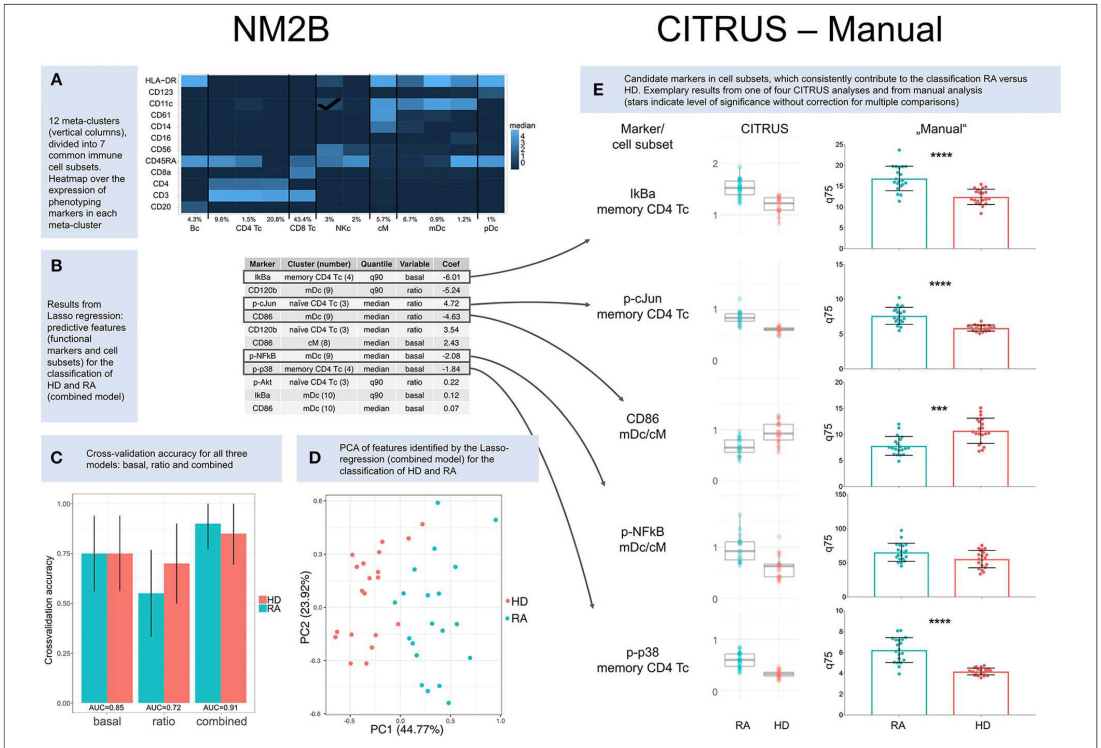



**FIGURE 1 |** Results. **(A)** Heatmap over the expression of phenotyping markers in 12 meta-clusters (columns). Meta-clusters were identified as B cells (Bc), $CD4^+$ T cells (CD4 Tc), $CD8^+$ T cells (CD8 Tc), natural killer cells (NKc), classical monocytes (cM), myeloid dendritic cells (mDc), and plasmacytoid dendritic cells (pDc). Relative abundance is given for each cell subset in percent. **(B)** Results from Lasso regression: predictive features (functional markers and cell subsets) and their contribution to the classification of healthy donors (HD) and RA patients (RA). Only nonzero coefficients are shown. Coefficients for CD86 and p-cJun are based on ratios and therefore inverted compared to CITRUS and manual comparisons of basal marker expression. **(C)** Cross-validation accuracy for all three NM2B analyses ("basal," "ratio," and "combined"), with area-under-the-curve (AUC) values for ROC analysis. **(D)** Principle component analysis (PCA) of features identified by Lasso-regression (combined model) for the classification of RA patients and healthy donors. **(E)** Cluster identification, characterization, and regression algorithm (CITRUS) and non-parametric testing (Manual). CITRUS results are presented in boxplots, as provided by the algorithm. Results from manual analysis are presented in scatter dot plots. Medians (CITRUS) and 75th percentiles (manual) are plotted for each RA patient (blue) and healthy donor (red); median, and upper and lower quartile. Asterisks indicate level of significance without correction for multiple comparisons (***$p < 0.001$, ****$p < 0.0001$).

Compared to the NM2B and manual analysis, there were slight differences in the weighting of cell subsets for p-cJun (**Figure 1E**). In four CITRUS analyses, memory $CD4^+$ T cells were the primary cell subset of interest for p-cJun, whereas automated analysis pointed to naïve $CD4^+$ T cells as the most significant cell subset. In manual analysis, p-cJun expression was significantly different between HD and RA in both naïve and memory CD4 Tc.

Exemplary CITRUS results and cross validation model error rates can be found in **Supplementary Figures 11, 12**.

## Pipeline 3: Manual Analysis

The manual analysis generally confirmed results from regression data modeling (**Figure 1E**). For the p-NFkB, regression tools had suggested significant differences in myeloid dendritic cell and classical monocyte subsets. However, in non-parametric testing, p-NFkB expression was not significantly different in myeloid dendritic cell and classical monocyte subsets of healthy donors vs. RA patients, but there were differences in p-NFkB in memory $CD4^+$ T cells. Complete results from manual analysis are shown in **Supplementary Figures 13–15**.

In summary, and based on all three analysis pipelines, we suggest that the phenotyping markers CD4, CD45RA, and CD11c (to identify CD4 naïve and memory $CD4^+$ T cells, and myeloid dendritic cells) and the functional markers p-p38, IkBa, p-cJun, p-NFkB, and CD86 may be candidate markers for a simplified setup, e.g., in confirmatory studies by flow cytometry.

## DISCUSSION

In this study, comprehensive investigation of signaling patterns in unstimulated and TNF-stimulated immune cells by mass

cytometry revealed cell type specific differences in RA patients compared to healthy donors of the same gender and with similar ages. Applying predictive regression models, we found that the basal expression of p-p38 and IkBa in memory $CD4^+$ T cells, p-cJun in naïve, and memory $CD4^+$ T cells, and p-NFkB and CD86 in myeloid dendritic cells and classical monocytes differentiated between healthy donors and RA patients. We want to emphasize the explorative character of our study and the role of mass cytometry in this setting. Mass cytometry and related analysis tools are currently not used in routine clinical practice. However, in this study we suggest a smaller set of markers for the distinction between HD and RA. These markers could make our approach applicable and feasible in future research e.g., on a flow cytometry platform. Our data indicate that phenotyping markers CD4, CD45RA, CD11c for the identification of $CD4^+$ T cell subsets and myeloid dendritic cells, and p-p38, IkBa, p-cJun, p-NFkB, and CD86 as relevant functional markers could be used to analyze unstimulated PBMCs by flow cytometry for diagnosis and stratification in RA.

Studies of signaling pathways in arthritis are often limited to one or two distinct cell subsets and a few functional markers and are frequently carried out in animal arthritis models. Comprehensive investigation of signaling in immune cell subsets in patients and healthy individuals has been challenging due to technical limitations. High-dimensional mass cytometry can fill a gap as it enables the simultaneous investigation of many markers in millions of heterogeneous cells with a single-cell resolution. Our study utilized a total of 34 channels (including barcoding and beads) and only partially exploited the potential of the technology.

In an analysis of a single RA patient and one healthy donor Nair et al. demonstrated that a complex mass cytometry setup distinguished between health and disease and was able to detect changes after TNF inhibitor treatment (12). Due to the illustrational character of their study, differences in signaling were not quantified, but both p-p38 and p-NFkB were differentiating markers in several cell subsets. Their data pointed to granulocytes as a cell population altered by TNF pathway activation. In support of this, another study had previously shown that granulocytes express high levels of TNF receptors (21). Unfortunately, our study did not include granulocytes as we studied PBMCs. PBMCs were selected to provide a detailed insight into non-granulocyte white blood cell populations, allowing for simultaneous stimulation of cells from the entire cohort under standardized conditions after cryo-preservation.

The use of cryo-preserved PBMCs introduces several potential contributors to variation, and deprives cells from their individual surroundings by the removal of plasma (22–26). We reduced variation through stringent use of standard operating procedures for the handling of live cells from the time of collection to cryo-preservation to resting and stimulation. Moreover, the experimental steps were conducted simultaneously on cells from all donors whenever feasible. However, for future study we would recommend the use of peripheral blood leukocytes with immediate fixation after sample acquisition from patient and donor.

Galligan et al. performed a phospho-flow analysis on PBMCs on a less homogeneous RA population than our cohort. The Galligan et al. cohort included RA patients at different disease stages treated with different medications and patients with osteoarthritis and healthy donors (11). In agreement with our results, they found elevated levels of several phospho-epitopes in $CD4^+$ T cell subsets in RA patients compared to healthy donors, and, to a lesser degree, to osteoarthritis patients. Interestingly, there were not significant differences in p38 phosphorylation levels between RA patients and healthy controls in the Galligan et al. study. Unfortunately, markers for the canonical NFkB-signaling pathway were not included.

To identify differences between "healthy" and "sick" representative cohorts of both groups are required. However, the number of simultaneously applicable barcodes, parallel handling of all samples, read-out time on the mass cytometer, and analysis of multi-dimensional data on millions of events set currently limits on cohort sizes. Based on a total of 40 individuals, our study is primarily of explorative character. Our cohorts were sex-matched. We aimed to achieve an age match between patients and healthy donors, although immune status has been shown to be rather stable over time in healthy adults (27). Our experimental setup allowed for a high degree of simultaneous analysis, running 80 samples (20HD+20RA in two conditions) at the same time. For future studies with more samples, it is important to assure that results are robust across different cytometry runs, e.g., through the use of a reference sample.

We only included newly diagnosed patients, in whom disease-related immune status was unaffected by immune-modulatory or immune-suppressive treatment with the exception of low-to-moderate dosages of prednisolone in five of the 20 patients. While prednisolone treatment may introduce an unwanted heterogeneity, this reflects the real-life situation at rheumatology outpatient clinics, with some patients being referred after pre-treatment. In a sub-group analysis with CITRUS we couldn't identify factors that differentiated prednisolone-treated from prednisolone-naïve patients.

RA is an inflammatory condition, and untreated patients are expected to express signs of inflammation on a cellular level compared to healthy donors. In our RA cohort, 15 patients had elevated levels of CRP, including all five patients treated with prednisolone. The higher levels of markers known to be involved in inflammatory signaling pathways, such as the canonical NFkB and the MAPK signaling pathways, in patients compared to healthy donors in this cohort was, therefore, not a surprise. We did not include patients with different inflammatory conditions in our study, hence the specificity of the observed signaling signatures for RA compared to other inflammatory conditions is not known. For future studies, cohorts with other TNF-driven conditions such as e.g., inflammatory bowel diseases should be added as disease controls.

That CD86 was consistently expressed to a lesser degree on classical monocytes and myeloid dendritic cells of RA patients compared to healthy donors is likely relevant to the pathogenesis of RA. CD86 is highly expressed on antigen-presenting cells in synovial fluid and synovia of RA patients, whereas CD28, the T

cell counterpart of CD86, is expressed at lower levels in patients with active RA compared to healthy donors (28).

In conclusion, this study provided insight into TNF-mediated signaling patterns, which are distinct for RA patients compared to healthy individuals. A comprehensive understanding of signaling signatures may facilitate more accurate diagnosis, better stratification of patients to guide treatment decisions, and the identification of candidate treatment targets in RA patients.

## DATA AVAILABILITY

The datasets for this study (normalized, concatenated, debarcoded, and after removal of events compromised by injector clogging can be found in the Flow Repository https://flowrepository.org/, experiment-ID FR-FCM-Z24N.

## ETHICS STATEMENT

All donors and patients gave written informed consent for inclusion into the Norwegian Arthritis Registry (NorArtritt) and the Research Biobank for Rheumatic Diseases in Western Norway (approval REK 2012/1689). Utilization of registry data and biobank material for this study was approved by the Regional Ethics Committee (approval REK 2014/317).

## AUTHOR'S NOTE

A summary of this work was presented in poster form at Cyto2018 in Prague, abstract B13 156.

## AUTHOR CONTRIBUTIONS

All authors have contributed to manuscript review. LB has contributed with experiment design, collection of samples and clinical data, cytokine and antibody titration, all laboratory work, data analysis pipeline 1 and 2, manuscript writing. S-EG has contributed to cytokine and antibody titrations, followed critically through especially the data analysis part of this study. NB and MB have contributed with data analysis pipeline 3 and manuscript writing. GB has contributed with laboratory work, ciritical follow-up through all stages of this study. AS has contributed with initial experiments for TNF stimulation and background work on TNF receptors/TNF receptor antibodies. CG has contributed to experiment design and facilitated collection of samples and clinical data. CV has contributed to experiment design and facilitated sample collection, laboratory work, and mass cytometry experiments. SG has contributed to experiment design and critical supervision throughout all parts of this study.

## FUNDING

This work was supported by the Center for Personalized Immunotherapy/the Western Norwegian Health authorities, research grant HV340029.

## ACKNOWLEDGMENTS

The authors want to thank all patients and healthy donors for their contributions to this work.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2019.01488/full#supplementary-material

## REFERENCES

1. Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. *Lancet.* (2009) 373:659–72. doi: 10.1016/s0140-6736(09)60008-8
2. Scott DL, Wolfe F, Huizinga TWJ. Rheumatoid arthritis. *Lancet.* (2010) 376:1094–108. doi: 10.1016/s0140-6736(10)60826-4
3. Bendtzen K. Personalized medicine: theranostics (therapeutics diagnostics) essential for rational use of tumor necrosis factor-alpha antagonists. *Discov Med.* (2013) 15:201–11.
4. Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez-Olivo MA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. *Can Med Assoc J.* (2009) 181:787–96. doi: 10.1503/cmaj.091391
5. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, III, et al. 2010 rheumatoid arthritis classification criteria: an american college of rheumatology/european league against rheumatism collaborative initiative. *Ann Rheum Diseases.* (2010) 69:1580–8. doi: 10.1136/ard.2010.138461
6. Bader LI, Solberg SM, Kaada SH, Bolstad N, Warren DJ, Gavasso S, et al. Assays for infliximab drug levels and antibodies: a matter of scales and categories. *Scand J Immunol.* (2017) 86:165–70. doi: 10.1111/sji.12572
7. Steenholdt C, Brynskov J, Thomsen OO, Munck LK, Fallingborg J, Christensen LA, et al. Individualised therapy is more cost-effective than dose intensification in patients with Crohn's disease who lose response to anti-TNF treatment: a randomised, controlled trial. *Gut.* (2013). doi: 10.1136/gutjnl-2013-305279
8. Cuppen BV, Welsing PM, Sprengers JJ, Bijlsma JW, Marijnissen AC, van Laar JM, et al. Personalized biological treatment for rheumatoid arthritis: a systematic review with a focus on clinical applicability. *Rheumatology.* (2016) 55:826–39. doi: 10.1093/rheumatology/kev421
9. Ermann J, Rao DA, Teslovich NC, Brenner MB, Raychaudhuri S. Immune cell profiling to guide therapeutic decisions in rheumatic diseases. *Nat Rev Rheumatol.* (2015) 11:541–51. doi: 10.1038/nrrheum.2015.71
10. Gavasso S, Gjertsen B, Anderssen E, Myhr K, Vedeler C. Immunogenic effects of recombinant interferon-beta therapy disrupt the JAK/STAT pathway in primary immune cells from patients with multiple sclerosis. *Multiple Sclerosis.* (2012) 18:1116–24. doi: 10.1177/1352458511434066
11. Galligan CL, Siebert JC, Siminovitch KA, Keystone EC, Bykerk V, Perez OD, et al. Multiparameter phospho-flow analysis of lymphocytes in early rheumatoid arthritis: implications for diagnosis and monitoring drug therapy. *PLoS ONE.* (2009) 4:e6703. doi: 10.1371/journal.pone.0006703

12. Nair N, Mei HE, Chen SY, Hale M, Nolan GP, Maecker HT, et al. Mass cytometry as a platform for the discovery of cellular biomarkers to guide effective rheumatic disease therapy. *Arthritis Res Ther*. (2015) 17:127. doi: 10.1186%2Fs13075-015-0644-z
13. Ornatsky O, Baranov VI, Bandura DR, Tanner SD, Dick J. Multiple cellular antigen detection by ICP-MS. *J Immunol Methods*. (2006) 308:68–76. doi: 10.1016/j.jim.2005.09.020
14. Ornatsky O, Bandura D, Baranov V, Nitz M, Winnik MA, Tanner S. Highly multiparametric analysis by mass cytometry. *J Immunol Methods*. (2010) 361:1–20. doi: 10.1016/j.jim.2010.07.002
15. Spitzer MH, Nolan GP. Mass cytometry: single cells, many features. *Cell*. (2016) 165:780–91. doi: 10.1016/j.cell.2016.04.019
16. Bagwell B. *A New Analytic Approach for Live Singlet Identification*. (2017). Available online at: https://www.fluidigm.com/articles/mass-cyto-summit-videos?mkt__tok=eyJpIjoiT0RVeU5ERTBNbU5oWW1GbSIsInQiOiJ4U3NDdUZjbVoyRCt6UThrWXg4b3M3RUNick5ZK0lsekJZQzU3VFZyU0QwYW5UaHdaQUxwbDh3ZHJzbTRcLzg3MlhIWkh5bEFDMUdWSFk0NWpiTmhueGVXQ3IzQzI4K0Q5a015cGM1bFdrVHl2elpNY2JGUERqb2I3MEk0RW9FemgifQ%3D%3D#bagwell (accessed June 09, 2017).
17. Fluidigm. *Maxpar Human Immune Monitoring Panel Kit Validation Studies*. (2018).
18. Kotecha N, Krutzik PO, Irish JM. Web-based analysis and publication of flow cytometry experiments. *Curr Protocols Cytometry*. (2010) Chapter 10:Unit10.7. doi: 10.1002/0471142956.cy1017s53
19. Bruggner RV, Bodenmiller B, Dill DL, Tibshirani RJ, Nolan GP. Automated identification of stratifying signatures in cellular subpopulations. *Proc Natl Acad Sci USA*. (2014) 111:E2770–7. doi: 10.1073/pnas.1408792111
20. Amir el AD, Davis KL, Tadmor MD, Simonds EF, Levine JH, Bendall SC, et al. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nat Biotechnol*. (2013) 31:545–52. doi: 10.1038/nbt.2594
21. Liu C, Tang J. Expression levels of tumor necrosis factor-α and the corresponding receptors are correlated with trauma severity. *Oncol Lett*. (2014) 8:2747–51. doi: 10.3892/ol.2014.2575
22. Corkum CP, Ings DP, Burgess C, Karwowska S, Kroll W, Michalak TI. Immune cell subsets and their gene expression profiles from human PBMC isolated by Vacutainer Cell Preparation Tube (CPT) and standard density gradient. *BMC Immunol*. (2015) 16:48. doi: 10.1186/s12865-015-0113-0
23. Grievink HW, Luisman T, Kluft C, Moerland M, Malone KE. Comparison of three isolation techniques for human peripheral blood mononuclear cells: cell recovery and viability, population composition, and cell functionality. *Biopreserv Biobank*. (2016) 14:410–5. doi: 10.1089/bio.2015.0104
24. Kadic E, Moniz RJ, Huo Y, Chi A, Kariv I. Effect of cryopreservation on delineation of immune cell subpopulations in tumor specimens as determinated by multiparametric single cell mass cytometry analysis. *BMC Immunol*. (2017) 18:6. doi: 10.1186/s12865-017-0192-1
25. Lemieux J, Jobin C, Simard C, Neron S. A global look into human T cell subsets before and after cryopreservation using multiparametric flow cytometry and two-dimensional visualization analysis. *J Immunol Methods*. (2016) 434:73–82. doi: 10.1016/j.jim.2016.04.010
26. Rahmanian N, Bozorgmehr M, Torabi M, Akbari A, Zarnani AH. Cell separation: potentials and pitfalls. *Preparat Biochem Biotechnol*. (2017) 47:38–51. doi: 10.1080/10826068.2016.1163579
27. Brodin P, Davis MM. Human immune system variation. *Nat Rev Immunol*. (2017) 17:21–9. doi: 10.1038/nri.2016.125
28. Sfikakis PP, Via CS. Expression of CD28, CTLA4, CD80, and CD86 molecules in patients with autoimmune rheumatic diseases: implications for immunotherapy. *Clin Immunol Immunopathol*. (1997) 83:195–8.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# Contents

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS


uib.no

ISBN: 9788230856031 (print)
9788230852309 (PDF)