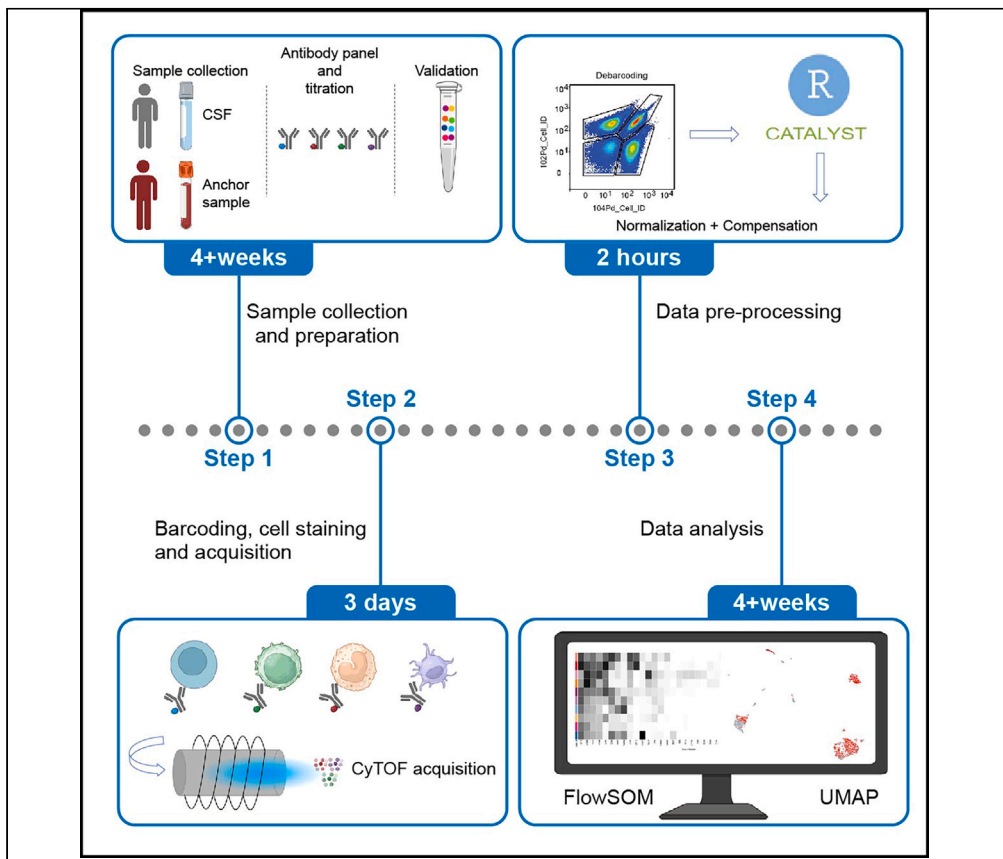


Protocol

Protocol to characterize immune cell subpopulations in cerebrospinal fluid of patients with neuroinflammatory diseases using mass cytometry



Gerardina Gallaccio,
 Meng Wang,
 Stephan
 Schlickeiser, Desiree
 Kunkel, Chotima
 Böttcher, Camila
 Fernández-Zapata

chotima.boettcher@
 charite.de (C.B.)
 maria.fernandez@charite.
 de (C.F.-Z.)

Highlights

Protocol to perform high-dimensional immune profiling of CSF cells using CyTOF

Instructions for sample collection, storage, staining, and measurement

Instructions for CyTOF data pre-processing and downstream analysis using R

Phenotypic and compositional changes of immune cells in cerebrospinal fluid (CSF) can be used as biomarkers to help diagnose and track disease activity for neuroinflammatory and neurodegenerative diseases. Here, we present a workflow to perform high-dimensional immune profiling at single-cell resolution using cytometry by time-of-flight (CyTOF) on cells isolated from the CSF of patients with neuroinflammation. We describe steps for sample collection and preparation, barcoding to allow for multiplexing, and downstream data analysis using R.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Gallaccio et al., STAR
 Protocols 5, 103038
 June 21, 2024 © 2024 The
 Author(s). Published by
 Elsevier Inc.
[https://doi.org/10.1016/
 j.xpro.2024.103038](https://doi.org/10.1016/j.xpro.2024.103038)



Protocol

Protocol to characterize immune cell subpopulations in cerebrospinal fluid of patients with neuroinflammatory diseases using mass cytometry

Gerardina Gallaccio,^{1,2,3,6} Meng Wang,^{1,2,3} Stephan Schlickeiser,⁴ Desiree Kunkel,⁵ Chotima Böttcher,^{1,2,3,6,*} and Camila Fernández-Zapata^{1,2,3,6,7,*}

¹Experimental and Clinical Research Center, A Cooperation between the Max Delbrück Center for Molecular Medicine in the Helmholtz Association and Charité-Universitätsmedizin Berlin, Berlin, Germany

²Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin, Germany

³Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

⁴Institute of Medical Immunology, BIH-Center for Regenerative Therapies, Charité-Universitätsmedizin Berlin, and Berlin Institute of Health Berlin, Berlin, Germany

⁵Berlin Institute of Health at Charité - Universitätsmedizin Berlin, Flow & Mass Cytometry Core Facility, Berlin, Germany

⁶Technical contact

⁷Lead contact

*Correspondence: chotima.boettcher@charite.de (C.B.), maria.fernandez@charite.de (C.F.-Z.)
<https://doi.org/10.1016/j.xpro.2024.103038>

SUMMARY

Phenotypic and compositional changes of immune cells in cerebrospinal fluid (CSF) can be used as biomarkers to help diagnose and track disease activity for neuroinflammatory and neurodegenerative diseases. Here, we present a workflow to perform high-dimensional immune profiling at single-cell resolution using cytometry by time-of-flight (CyTOF) on cells isolated from the CSF of patients with neuroinflammation. We describe steps for sample collection and preparation, barcoding to allow for multiplexing, and downstream data analysis using R. For complete details on the use and execution of this protocol, please refer to Fernández-Zapata et al.¹

BEFORE YOU BEGIN

Neuroinflammation plays a crucial role in chronic and acute central nervous system (CNS) disorders, with multiple sclerosis (MS) being the most prevalent immune-related CNS disorder. We suggest a workflow for in-depth analysis of immune cells in the cerebrospinal fluid (CSF) of patients with neuroinflammation, applicable for immune-cell profiling in the CSF of individuals with MS or other neuroinflammatory diseases.

In MS the inflammatory response is orchestrated by infiltrating autoreactive lymphocytes which in combination with co-migrating blood-derived macrophages and activated microglia form inflammatory foci in the white matter of the CNS.² The CNS has traditionally been regarded as a site of immune privilege, protected by specialized barriers, including the blood-cerebrospinal fluid (CSF) barrier and CNS-CSF barrier. However, it is neither immunologically separated from the peripheral immune system, nor inert to damage signals from the peripheral organs³. Under homeostatic conditions, entry of immune cells from the periphery into the CNS parenchyma is under tight control, with little evidence for patrolling immune cells other than CNS-resident microglia and few tissue-resident CD8⁺ memory T cells instigated by previous infections. Nevertheless, indirect surveillance of the CNS exists, based on transport of parenchymal antigens together with interstitial fluid into



CSF.⁴ The CSF hosts a subset of blood immune cells predominantly composed of memory T cells, but also B cells, monocytes, natural killer (NK) cells, unconventional T cells and antigen-presenting cells (APCs). Routine CSF collection is integral for diagnosing various CNS disorders, offering a more practical alternative to assess the CNS. Thus immunophenotyping of CSF cells is feasible and can provide better insights in immune-driven/mediated pathophysiology of many neurological disorders including neuroinflammation.⁵ Moreover, unraveling cellular biomarkers can potentially be used as a diagnostic tool and as a measure of disease activity, thereby facilitating more personalized treatment approaches and enhancing our comprehension of the underlying factors contributing to diseases heterogeneity. Taking advantage of the minimal spillover between channels and no interference by auto-fluorescence, Cytometry by Time-of-Flight (CyTOF) is a powerful tool for comprehensive characterization of multiple immune cell populations in different body compartments. It allows a simultaneous evaluation of over 40 protein markers at the single cell level. The combination of a comprehensive array of protein markers and unsupervised data analysis provides a powerful strategy for profiling the heterogeneity of human immune cells in health and disease.^{6,7} However, an implementation of CyTOF-based immune phenotyping studies in neuroimmunology research is limited by complex experimental workflows and the variation in samples composition between batches. Therefore, it is important to standardize and streamline the experimental and analysis workflows starting from sample preparation, especially in longitudinal studies with a large cohort of patients. We describe here a standardized and streamlined workflow from sample collection and processing to data acquisition and analysis of small numbers of immune cells in CSF. Historically, it has been technically challenging to perform immune profiling of CSF cells due to small cell numbers (~5,000–15,000 cells/mL) and vulnerability of CSF immune cells (which limits possibility of cryopreservation), as well as availability of the patient CSF samples. Furthermore, our protocol can be also applied for isolated single cells from other tissues such as brain or from peripheral blood.

Of note, in this protocol, we do not aim to describe in detail about antibody panel design and titration, which have previously been well described by Thrash et al., STAR protocols (2020).⁸

Institutional permissions

Since the protocol involves human biomaterials, an approval of Ethics Committee is necessary before starting the sample collection and/or analysis of biomaterials. The study was registered and approved by the Ethics Committee of Charité – Universitätsmedizin Berlin. All participants provided written informed consent before any study-related procedures were undertaken. The study will be performed according to the Declaration of Helsinki and to the relevant ethical guidelines for research in humans.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Human HLA-DR(L243)-141Pr (1:100)	BioLegend*	Cat#307651
Human CD19(HIB19)-142Nd (1:100)	Standard BioTools	Cat#3142001B
Human cPARP(F21-852)-143Nd (1:100)	Standard BioTools	Cat#3143011A
Human CD69(FN50)-144Nd (1:100)	Standard BioTools	Cat#3144018B
Human CD4(RPA-T4)-145Nd (1:100)	Standard BioTools	Cat#3145001B
Human CD64(10.1)-146Nd (1:100)	Standard BioTools	Cat#3146006B
Human CD226(REA1040)-147Sm (1:100)	Miltenyi*	Cat#130-126-485
Human CD16(3G8)-148Nd (1:50)	Standard BioTools	Cat#3148004B
Human CD56/NCAM(NCAM16.2)-149Sm (1:100)	Standard BioTools	Cat#3149021B
Human MIP1beta/CCL4(D21-1351)-150Nd (1:50)	Standard BioTools	Cat#3150004B
ICOS/CD278 (C398.4A)-151Eu (1:50)	Standard BioTools	Cat#3151020B
Human CD66b(80H3)-152Sm (1:100)	Standard BioTools	Cat#3152011B

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human CD68(Y1/82A)-153Eu (1:100)	BioLegend*	Cat#333802
Human CD3(UCHT1)-154Sm (1:200)	Standard BioTools	Cat#3154003B
Human CD11c(Bu15)-155Gd (1:200)	BioLegend*	Cat#337221
Human IL-6(MQ2-13AS)-156Gd (1:100)	Standard BioTools	Cat#3156011B
Human CCR4/CD194(L291H4)-158Gd(1:200)	Standard BioTools	Cat#3158032A
Human TIGIT(MBSA43)-159Tb (1:50)	Standard BioTools	Cat#3159038B
Human CD14(RMO52)-160Gd (1:100)	Standard BioTools	Cat#3160006B
Human CTLA4/CD152(14D3)-161Dy (1:100)	Standard BioTools	Cat#3161004B
Human CD8(RPA-T8)-162Dy (1:50)	Standard BioTools	Cat#3162015B
Human CRTH2/CD294 (BM16)-163Dy (1:100)	Standard BioTools	Cat#3163003B
Human CD95(DX2)-164Dy (1:50)	Standard BioTools	Cat#3164008B
Human LAG-3/CD223(11C3C65)-165Ho (1:100)	Standard BioTools	Cat#3165037B
Human CD141(M80)-166Er (1:100)	Standard BioTools	Cat#3166017B
Human CCR7/CD197(G043H7)-167Er (1:100)	Standard BioTools	Cat#3167009A
Human CD206(15-2)-168Er (1:100)	BioLegend*	Cat#321102
Human CD33(WM53)-169Tm (1:100)	Standard BioTools	Cat#3169010B
Anti-biotin-(1D4-C5)-170Er (1:100)	Standard BioTools	Cat#3170003B
Human CD161/KLRB1(HP-3G10)-171Yb (1:100)	BioLegend*	Cat#339919
Human OPN/SPP1(polyclonal)-172Yb (1:200)	LSBio*	Cat#LS-C99283
Human CXCR4/CD184(12G5)-173Yb (1:200)	Standard BioTools	Cat#3173001B
Human IL-1 beta(CRM56)-174Yb (1:50)	eBioscience*	Cat#14-7018-85
Human TNF(Mab11)-175Lu (1:25)	DVS Sciences/Standard BioTools	Cat#3175023B
Human CD127/IL-17Ra(A019D5)-176Yb (1:200)	DVS Sciences/Standard BioTools	Cat#3176004B
Human CD47(CC2C6)-209Bi (1:200)	Standard BioTools	Cat#3209004B
Chemicals, peptides, and recombinant proteins		
Cell-ID 20-plex barcoding kit (https://fluidigm.my.salesforce.com/sfc/p/#7000000_09DAw/a/4u0000019iFV/_MMkFk_8Oes9hcoYhMMr6Drg2mgwF69GVV0UAFUgeQc)	Standard BioTools	Cat# PN 201060
Maxpar X8 multimetal labeling kit-40Rxn (https://fluidigm.my.site.com/Storefront/Cytometry/ConsumablesandReagentsCytometry/MaxparAntibodyLabelingKits/Maxpar%C2%AE%20X8%20Multimetal%20Labeling%20Kit%E2%80%9440%20Rxn?ccl=en_US)	Standard BioTools	Cat#201300
Smart Tube PROT1 (https://www.smarttubeinc.com/)	Smart Tube Inc.	Cat#501351689
FC receptor blocking reagent	Milteny Biotec	Cat# 5170126102
Bovine serum albumin (BSA)	Carl Roth	Cat#T8844.2
RPMI medium: RPMI-1640, 10% FBS	Gibco	Cat#31870-025
PBS	Gibco	Cat#14200-067
Methanol-free formaldehyde (FA) solution	Thermo Fisher Scientific	Cat#28906
Universal nuclease	Thermo Fisher	Cat##88701
Software and algorithms		
R v4.3.0 (Already tomorrow)	CRAN	https://www.r-project.org/
R studio v2023.03.1-446	Posit	https://posit.co/download/rstudio-desktop/
Bioconductor	Bioconductor- open-source software for bioinformatics	https://www.bioconductor.org/
CATALYST	CATALYST: cytometry data analysis tools.	https://github.com/He lenaLC/CATALYST .
BioManager	Bioconductor- open-source software for bioinformatics	https://www.bioconductor.org/install/
flowCore	flowCore: basic structures for flow cytometry data	https://www.bioconductor.org/packages/release/bioc/html/flowCore.html
flowWorkspace	Infrastructure for representing and interacting with gated and ungated cytometry data sets.	https://www.bioconductor.org/packages/release/bioc/html/flowWorkspace.html
CytoML	A GatingML interface for cross platform cytometry data sharing	https://www.bioconductor.org/packages/release/bioc/html/Cyto ML.html

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
flowFP	Fingerprinting for flow cytometry	https://www.bioconductor.org/packages/release/bioc/html/flowFP.html
diffcyt	Differential discovery in high-dimensional cytometry via high-resolution clustering	https://github.com/lmweber/diffcyt
scater	Single-Cell analysis toolkit for gene expression data in R	https://www.bioconductor.org/packages/release/bioc/html/scater.html
ggplot2	The R Graph Gallery	https://r-graph-gallery.com/ggplot2-package.html
CyTOF software v7.1	Fluidigm	https://go.fluidigm.com/cytofsw/v7
FlowJo v10.9	FlowJo	https://www.flowjo.com/solutions/flowjo/downloads

Other

5 mL serological pipettes	Eppendorf	Cat#0030127714
1.5 mL Eppendorf tube	Eppendorf	Cat#50-809-150
Falcon tubes	A Corning Brand	Cat#352096
Water bath at 37°C	Memmert	Cat#WNB45
Centrifuge	Eppendorf	Cat #5418R
Filter tips	Starstedt	Cat#70.3050255

* These antibodies were conjugated in house using the Maxpar X8 antibody labeling protocol (see step 3)

STEP-BY-STEP METHOD DETAILS**Sample collection and storage**

⌚ **Timing:** 4+ weeks

In this section we highlight all the good practices' phases to perform CSF collection and storage, including collection of anchor samples used to normalize signal between measurements.

1. Prepare the anchor sample.

Note: An anchor sample is peripheral blood mononuclear cells (PBMCs) used as internal reference across different measurements/batches to facilitate the signal normalization.⁶ However, cell types other than PBMCs can also be used as an anchor sample but they should properly express all the markers of desired panel(s). Number of anchor samples to be prepared depends on number of planned batches. One anchor sample will be added to each batch of pooled samples.

- Mix the isolated PBMCs (1×10^6) with 250 μ L of 10% bovine serum albumin (BSA) (in PBS).
- Add 350 μ L of proteomic stabilizer (PROT1) buffer, gently mix and incubate at room temperature (RT) for 12 min.
- Immediately store at -80°C .

⚠ CRITICAL: Centrifugation speed for living cells should not exceed $300 \times g$. This step is critical; be careful not to spin down too fast during the isolation of PBMC from whole blood. This precaution helps overcome high cell loss caused by increased cell lysis and membrane deformation.

⏸ Pause point: Samples can be stored at -80°C for up to 4 years. Proteomic Stabilizer PROT1 can be stored at room temperature for up to 5 years. It is recommended to re-test the activity of Proteomic Stabilizer PROT1 every two years to ensure there has been no loss of activity. Alternatively, for a long-term cell and marker preservation protocol, such as the one illustrated here, the Phosphoflow Fix and Lyse (BD Biosciences, Franklin Lakes, NJ) buffer can be used.

However, careful selection of suitable buffers and compatible antibodies is absolutely required since problems in the detection of some markers and population are more affected in one or the other fixation method.⁹

2. CSF sample collection and processing (for 3 mL CSF).

△ **CRITICAL:** CSF samples must be kept on ice and should be processed (i.e., cell isolation and aliquoting) within maximum of one hour after lumbar puncture. CSF cell pellet is commonly invisible (due to low cell numbers, commonly about 5,000–10,000 cells/ml), thus sample processing must be performed with care to avoid disturbing the cell pellet (see also troubleshooting [problem 1](#)).

- Place 3 mL CSF in 15 mL Falcon polypropylene (PP) tube.
- Centrifuge at $300 \times g$ (4°C) for 10 min.
- Carefully take out the supernatant. To avoid disturbing the cell pellet, about 100 µL of CSF is left in the tube. CSF supernatant may be aliquoted for e.g., proteomics or metabolomics analysis.
- Add 400 µL of 10% BSA (in PBS) to the CSF cell pellet, gently mix by pipetting up and down for about 4–5 times.
- Add 700 µL of PROT1 buffer, gently mix and incubate at RT for 12 min.
- Immediately store at –80°C.

△ **CRITICAL:** After the centrifugation (after b.) CSF samples must be clear and without any blood contamination (depicted by the presence of an erythrocyte pellet). CSF samples contaminated with blood cells should be excluded from the study (see the troubleshooting [problem 2](#)).

3. Antibody panel preparation.

Note: In-depth descriptions of the antibody titration and antibody panel validation steps have been extensively described in previous STAR protocol by Thrash et al.⁸ For the steps regarding the validation and conjugation of antibodies, comprehensive information can be accessed through the Standard Biotools website (<https://fluidigm.my.salesforce.com/sfc/p/#700000009DAw/a/4u0000019jXU/6wyoqHHEDHI 5D5e0cLOsylAsnfB0hdiCEprKHI9aFj8>).¹⁰

- Preparation of antibody cocktail and storage:

Note: For example, for five batches of pooled samples, a final volume of 500 µL of an antibody cocktail is prepared, consisting of a combination of all antibodies in each panel. Each antibody is diluted in staining buffer according to a validated dilution. The final antibody cocktail is subsequently divided into 5 aliquots (100 µL each) and stored at –80°C. Each pooled sample (i.e., one batch containing max. of 20 individual samples) will be stained with a 100 µL frozen antibody cocktail.

▣ **Pause point:** Regarding the storage conditions, the X8 polymer is moisture sensitive. The polymer tubes must be stored at –20°C with provided desiccant and sealed container and should be used only once and immediately after thawing to room temperature. Instead, the antibody cocktail can be stored for 6 months up to 1 year at –80°C without degraded signal.

Barcoding and staining

⌚ **Timing:** 2 days

This step describes sample barcoding and consequent staining with antibody cocktails.

4. Sample barcoding:

Note: For sample barcoding, we use the Cell-ID-20-plex Pd Barcoding Kit (Standard Biotools), containing a total of 20 different metal combinations, using a 6-choose-3 barcoding scheme (combinations of any 3 Palladium (Pd) isotopes of ^{102}Pd , ^{104}Pd , ^{105}Pd , ^{106}Pd , ^{108}Pd , or ^{110}Pd). Barcoding can also be used as a tool to identify cell-cell doublets. To minimize variability of inter-sample staining and acquisition, we pool all samples (max. of 20 samples) into one batch prior to staining with frozen antibody cocktail¹¹ as described below.

- a. Transfer a maximum of nineteen samples and one anchor sample (all PROT1-fixed samples) from -80°C on dry ice.
- b. Thaw samples at 4°C , until completely thawed.
- c. Transfer cells into a 15 mL Falcon PP tube, containing 10 mL cell staining buffer (CSB) (Standard Biotools).
- d. Centrifuge at $600 \times g$ (RT) for 5 min. Discard the supernatant.
- e. Gently resuspend cell pellet in 1 mL CSB and transfer cell suspension into a 1.5 mL Eppi. Centrifuge at $600 \times g$ (RT) for 5 min. Discard the supernatant.
- f. Gently resuspend cell pellet in 1 mL Benzonase medium (RPMI1640+FBS (10%) diluted 1:10000). Add additional 4 mL warmed Benzonase medium (37°C). Pipette gently up and down.
- g. Incubate at RT for 30 min.
- h. Wash twice with 1 mL CSB. Centrifuge at $600 \times g$ (RT) for 5 min. Discard the supernatant.
- i. Pool all twenty samples in a new 1.5 mL Eppi. Centrifuge $600 \times g$ (RT) for 5 min. Discard the supernatant.
- j. Gently resuspend the pellet in 20 μL of 1 mg/mL Beriglobin in CSB (Stock sol = 160 mg/mL, diluted 1:160 ratio in staining buffer). Incubate for 10 min at 4°C .
- k. Add 90 μL of thawed (frozen at -80°C) antibody cocktail, resuspend and incubate at 4°C for 30 min.
- l. Wash twice with 1 mL CSB and centrifuge at $600 \times g$ (4°C) for 5 min. Aspirate the supernatant carefully.
- m. For the fixation step, gently resuspend cell pellet in 500 μL of freshly prepared 2% methanol-free formaldehyde (FA) solution (Pierce, freshly diluted from stock solution in PBS) (max. of 1 Mio cells per 100 μL FA solution, rotate with a volume bigger than 500 μL) and incubate overnight at 4°C .

Note: Formaldehyde is commonly used as a fixative in biological and medical research to preserve cells or tissue samples. Its ability to cross-link proteins and nucleic acids helps to "fix" cellular structures, maintaining their integrity for further analysis or experimentation. Due to the high chemical reactivity, formaldehyde causes local irritation and acute and chronic toxicity after direct contact in target tissues. Considering the mode of action (MOA) formaldehyde is classified in Category 4 for carcinogens.¹²

- n. On day 2: Wash cell suspension once with 1 mL CSB. Centrifuge at $800 \times g$ (4°C) for 5 min and discard the supernatant.
- o. Thaw frozen (-80°C) intracellular antibody master mix on ice. Mix 50 μL of antibody cocktail with 50 μL of 2 \times Perm Buffer (eBioscience) to fix the cells.
- p. Resuspend the cell pellet in 90 μL of the diluted antibody cocktail. Incubate at RT for 30 min.
- q. Wash twice in 1 mL CSB, centrifuge at $800 \times g$ (4°C) for 5 min. Add 500 μL Iridium solution (diluted 1:1000 in PBS containing 2% FA). Incubate at RT for 20 min.
- r. Wash twice with 1 mL CSB. Centrifuge at $800 \times g$ (4°C) for 5 min. Discard the supernatant.
- s. Keep at 4°C in max. of 80 μL CSB at this point until ready for CyTOF measurement.

- t. Transfer cells in max. of 80 μ L CSB on the strip for washing with MilliQ H₂O using the Laminar Wash Mini-1000 system (CuriOX Biosystems). Settle down for 30 min, wash immediately before measurement (9 cycles, FR 5).
- u. CyTOF acquisition.

Acquisition

⌚ Timing: 1 day

In this section, we highlight all the good practices' phases to achieve optimal data acquisition with the CyTOF instrument.

5. Sample acquisition.

Note: For detailed information please refer to the related paper: Mass Cytometry, Methods and Protocols, Helen M. McGuire and Thomas M. Ashhurst, 2019.¹³ Before each use, a quality control of the instrument should be performed and documented for performance tracking. This quality control should at least include:

- a. Contamination check: before the instrument is tuned to the manufacturer's instructions, short preview with ultrapure water will show remaining metal contamination.
- b. Quality control: EQ four element beads from Fluidigm can be run for a defined period of time (e.g., 2 min) to control for yield and sensitivity, and can serve as a quality control before each experiment.
- c. Samples check: It is advisable to resuspend samples one by one and run samples with a large number of cells in aliquots of no more than 1 h runs (approx. one million cells), as the cells tend to disintegrate over time if they are kept in water, even if fixed well. It is important to adjust the cell concentration directly before running the sample.

Note: After sample acquisition, the data are digitized as a raw integrated mass data (IMD) file, which represents a matrix of ion counts for each selected mass channel for every push. The Fluidigm software then converts the IMD to a flow cytometry standard (FCS) file. The resulting FCS file contains total integrated ion counts for every selected channel for every event and can be analyzed using FlowJo, Cytobank, or other available third-party cytometry analysis software.

Note: Optimal sample concentration might differ depending on the source of the cells or cell type. Isolated cells from liquid biopsies (e.g., blood, liquor, urine) are less prone to form clumps than cells isolated from tissue and can be run at higher concentrations. There is a maximum number of events/second that is around 300–500 events/s.

⚠ CRITICAL: Good cell fixation is extremely important to allow cells to withstand the hypotonic stress associated with the final water washes prior to sample acquisition; all stained samples should be postfixed with freshly diluted 1–4% formaldehyde in PBS. Inadequate fixation will result in sample degradation, which may manifest as significant loss of cells during the water wash. Even in appropriately fixed samples, exposure to water for prolonged periods of time will cause cellular degradation.

Note: Cell signal intensity may decrease during acquisition as result of instrument performance or sample degradation. Changes in instrument performance can occur between samples and even within a single acquisition of sample. This can be due to gradual loss of detector sensitivity, or changes in plasma ionization efficiency. Fluidigm's EQ Four Element Calibration Beads are polymer beads that contain known standards of four elements at natural isotopic

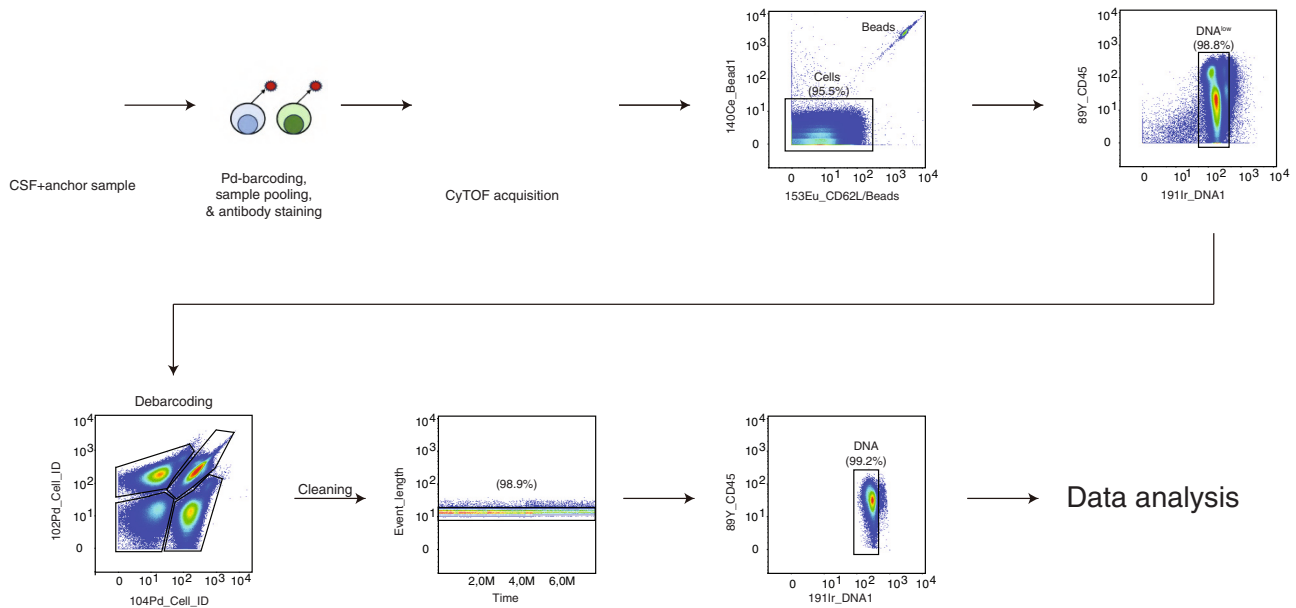


Figure 1. Representative plots of sample barcoding, staining and gating strategy

Samples are barcoded, pooled, and stained with a panel of metal-conjugated antibodies and acquired on the CyTOF instrument. Prior to the data analysis each individual sample is de-barcoded on FlowJo. Figure adapted with [BioRender.com](https://www.biorender.com).

abundance (cerium, europium, holmium, and lutetium). Several metrics can be evaluated using EQ bead-derived signals to track daily instrument performance and to monitor changes during sample acquisition using conventional cytometry analysis software or automated tools. While instrument performance can be tracked using EQ beads, cell-specific degradation cannot. Once a sample has been acquired and undergone initial quality control several steps must be taken to process the data for analysis. As mentioned previously, instrument performance can vary within a single sample. To address this issue, EQ bead normalization must be performed to account for this technical variability in order to better represent real biological differences between samples. These beads allow for monitoring of instrument performance and for normalization of signal intensity to account for fluctuations over time or variations between instruments.

Pause point: On Helios or CyTOF XT instrument EQ Six Element Calibration Beads (Standard BioTools) could also be used as an alternative of EQ Four Element Calibration Beads.

Data pre-processing

⌚ Timing: 4+ weeks

In this section we describe the data preprocessing steps including de-barcoding of individual samples, normalization of batch effect and compensation to correct for signal spillover.

6. Debarcoding.

Note: In this section we describe the gating strategy to identify individual samples (Figure 1). Raw data (.FCS files) are generated after CyTOF acquisition. Each raw data (.FCS file) comprises 20 barcoded samples. Using FlowJo Software, calibration beads will be removed from the file, cells are identified based on DNA signal (i.e., 192/193Ir). For debarcoding,

Boolean gating is used to deconvolute individual samples according to the barcode combination. All de-barcoded samples are then exported as individual FCS files for downstream data analysis.^{1,6}

7. Install tools/ software, packages and libraries.

Note: This protocol utilizes the R environment for statistical computing and data visualization. Depending on your operating system, download the necessary software from the CRAN repositories (<https://cran.r-project.org/>) and RStudio website. Install the required R tools accordingly.

After installing the latest software version, install all the R packages specific to CyTOF data analysis, by executing the following command on the RStudio console. Here we show the libraries used to execute our workflow. The key resources table lists the R packages used.

△ CRITICAL: It is highly recommended to verify that the user is running the data analysis in the latest version of the R environment, as previous version of the software may return some errors. Additionally, please note that R Tools dependencies may be required to access the necessary libraries in Windows software. The function `install.packages()` could return a “lib path error”. (See also [troubleshooting problem 3](#))

```
if (!require("BiocManager", quietly = TRUE))
install.packages
("BiocManager") BiocManager::install("CATALYST")
BiocManager::install("flowCore")
BiocManager::install
("flowWorkspace")
BiocManager::install("CytoML")
BiocManager::install("flowFP")
BiocManager::install("diffcyt")
BiocManager::install("scater")
```

8. Normalization.

Note: Due to availability, human neuroimmunology studies often involve sample collection over the course of months to years. To curate enough dataset for powerful statistical testing, it is necessary to process and run samples in multiple batches over a period of time.

A barcoding approach allows for multiple samples to be stained together in one tube, reducing the intra-barcode technical variability, and optimizing data acquisition speed and efficiency (resulting in decreased cell loss), as it constitutes a single sample run on the instrument. However, it is only possible to measure 20 samples per barcode set and so, multiple barcode sets (batches) are still required to address questions in robustly powered study designs. To improve integration of data from different batches, thus to minimize the batch effect, we must integrate a signal-normalization step in the data analysis workflow.

For detailed information, please refer to the related paper “Minimizing Batch Effects in Mass Cytometry Data, Schuyler et al. 2019”¹⁴

- a. Before starting the Normalization step, create a folder containing: The script “Normalization_R”; The ChannelsToAdjust_example.txt file containing a list of the channel names used in the experiment; The function BatchAdjust_R, which can be downloaded from <https://github.com/CUHIMSR/CytofBatchAdjust>. The compensated FCS files in a subfolder named “Files”.

b. Rename your FCS files:

The FCS files will be divided in groups of batches: eg., Batch1, Batch2, Batch3 etc. Each batch group must contain an anchor sample. Rename the FCS files by adding the word "BatchNumber_" with the corresponding number of the batch group they belong to and add the word "anchor" for every anchor sample.

▣▣ **Pause point:** Take a look at this explanatory example before proceeding.

You have a series of FCS files that belong to Batch 1 and 2. You could rename the file as follows. Sample_number_batchNumber_SampleName

Sample1_Batch1_013.fcs

Sample2_Batch1_014.fcs

Sample3_Batch1_015.fcs

Sample4_Batch2_033.fcs

Sample5_Batch2_034.fcs

Do the same for the anchor samples. You must have only one anchor sample for each Batch.

Number of anchor sample_number of batch_anchor

Sample01_Batch1_anchor.fcs

Sample02_Batch2_anchor.fcs

Sample03_Batch3_anchor.fcs

(Sample01, Sample02, Sample03 if they belong to Batch1, Batch2, Batch3)

c. Run the script.

```
# Change the directory accordingly!

# Note you must have a basedir with your original files ONLY and the function will create an empty
output folder called ``Out``.

# you need a .txt file with your channels to adjust, check the example and adjust to your own

# the batch key word and anchor key word are very important, you need to change your files names
accordingly

library(flowCore)

# Call the function
source("Normalization/BatchAdjust.R")

# Directory containing original
files filedir <- "Normalization/Files/"

# Normalize the files accordingly with the batches and anchor samples

# Chose the percentile method, 95th or 80th

BatchAdjust(basedir=filedir,
outdir="C:/Users/admin/Desktop/Normalization/Out", channelFiles="
Normalization/ChannelsToAdjust_example.txt, batchKeyword="Batch", anchorKeyword= "an-
chor", method="95th")
```

Note: Use the 95th percentile as the high end for our normalization target point to avoid outliers, and 80th percentile as the low end.

d. Obtain the output graphs:

Take a look at the pre- and post-normalization variance and verify that the signal intensity of markers in each channel is correctly normalized.

▮ **Pause point:** The normalization script provided here is highly sensitive and must function correctly for both Windows and Mac users. We strongly advise users to run the script exactly as shown here to prevent errors due to case sensitivity or typing mistake.

⚠ **CRITICAL:** You might encounter an error that prevents you from obtaining the normalized output files and the corresponding plots. This issue could be related to the name of the channel-markers listed in the text file. (See also troubleshooting [problem 4](#))

- e. Obtain the output of normalized FCS files.
Use the normalized FCS files to proceed with the Compensation step.

9. Compensation.

Note: This section will drive into the evaluation and compensation of signal spillover. Although signal spillover in CyTOF is minimal compared to fluorescent-based technologies, there is still signal crosstalk between channels that can interfere with the interpretation of the data. This spillover is mainly due to natural isotopic impurity ($m + 1$, $m + 2$, etc.) and oxidation of elements during measurement ($m + 16$). The spillover is correlated with the original signal in an approximately linear manner and can be corrected via a process called compensation.¹⁵ In parallel to multiplexed sample staining, single stains were generated by staining polystyrene antibody- capture beads (SS-beads). After staining, beads were pooled and run as a single sample in the mass cytometer. Each bead is assigned to a specific population based on the dominant signal, and the purity of the bead populations is further increased by automatically applying estimated sample-specific cutoffs. In a second step, the spillover matrix is calculated based on the spillover observed for single-stained populations.¹⁵ This workflow primarily relies on the usage of CATALYST packages¹⁶ (<https://github.com/HelenaLC/CATALYST>), which are necessary for performing CyTOF data analysis.

Before starting: Create a folder that contains the SS-beads FCS file, unzipped FCS files, and the script needed for the analysis.

- a. Load the libraries needed for this step.

```
library(CATALYST)
library(flowCore)
library(SingleCellExperiment)
```

Note: CATALYST performs compensation via a two-step approach comprising identification of single positive populations via single-cell debarcoding (SCD) of single-stained beads (or cells) and estimation of a spillover matrix (SM) from the populations identified, followed by compensation via multiplication of measurement intensities by its inverse, the compensation matrix (CM).

- b. Data organization.

Load the single stains data and make sure to have SS_Beads FCS file in the working directory. Data are organized into an object called SingleCellExperiment (SCE)¹⁶ which can be constructed from a directory housing a single or set of FCS files. FCS files are read into R with read.FCS function of the flowCore package and are represented as an object of class flowFrame.¹⁶

```
# Load the single-stained beads (SS_Beads) and address the parameters Single_stains <-
  ``SS_Beads_01.FCS``
# Read the single_stained beads as FCS file
```

```

ss_exp <- read.FCS(single_stains, transformation=FALSE, truncate_max_range=FALSE)

# Specify mass channels stained for and set as numeric
bc_ms <- as.numeric(gsub("[[:alpha:]]", "",
  apply(strsplit(parameters(ss_exp)$desc, "_"), '[', 1)))

# Check for NA values
bc_ms <- bc_ms[!is.na(bc_ms)]

bc_ms <- bc_ms[!(bc_ms %in% c(89, 113, 115, 140, 190, 191, 193, 195))]

```

c. Debarcoding.

Note: The debarcoding process commences by assigning a preliminary barcode ID to each event.

- i. assignPrelim function will return either a binary barcoding scheme or a vector of numeric masses as input, and accordingly assigns each event the appropriate row name or mass as ID.
- ii. Final assignment will be made by applyCutoffs function.
- iii. plotYields, shows the distribution of barcode separations and yields upon debarcoding as a function of separation cutoff.

```

# Prepare the data
re<-prepData(ss_exp)

# Assign the preliminary barcode ID
re <-assignPrelim(re, bc_ms, verbose = FALSE)

# Apply the cutoffs
re <-applyCutoffs(estCutoffs(re))
re <-estCutoffs(x=re)
sep_cutoffs<-re$sep_cutoffs

re <-applyCutoffs(x=re, sep_cutoffs = sep_cutoffs)

# Visualize the single stained bead deconvolution
plotYields(x=re, which=0)

```

d. Compensation.

Note: These steps are relevant to the compensation of FCS files.

- i. Extract the spillover matrix: The following functions, computeSpillmat and plotSpillmat, provided an estimation and visualization of the spillover matrix for channels intensities signal.

```

re<-computeSpillmat(re)

# Check the channels and metals
sm<-metadata(re)$spillover_matrix
chs<-channels(re)

ss_chs<-chs[rowData(re)$is_bc]

```

```
all (diag(sm[ss_chs, ss_chs]) == 1)
all (sm >= 0 & sm <= 1)

custom_isotope_list <- c(CATALYST::isotope_list, list(BCKG=190))

# Get the Spill matrix plot before the compensation of the datasets plotSpillmat(re,
isotope_list=custom_isotope_list)
```

Note: The SM is stored in the SCE object as well as the custom_isotope list.

- ii. compCytof function permits to compensate mass cytometry-based experiments using a provided spillover matrix.

```
# Use the ``flow`` method
re_c <- compCytof(re, sm, method = "flow", isotope_list=custom_isotope_list)
fs <- sce2fcs(re_c)

exp_dat <- exprs(fs)
set.seed(25)

exp_dat <- asinh(exp_dat[sample.int(nrow(exp_dat), 5000), c(7:50)]/5)

# Obtain the first scatter plot matrix
pairs(exp_dat, pch=".")
```

- iii. Select a random sample, for instance "sample1", within the dataset used.

```
# Random chosen sample: sample1 # Load the sample1
sample1 <- read.FCS("sample_01.fcs", transformation=FALSE, truncate_max_range=FALSE)

# Check the info stored in the sample1
sample1

# Address all the parameters to sample1 before performing the compensation
sce <- prepData(sample1)

sce <- assignPrelim(sce, bc_ms, verbose = FALSE)

# Look at the information stored in sample1 (desc function) and select the numbers corresponding to the right channels.
exp_dat <- exprs(sample1)

exp_dat <- asinh(exp_dat[sample.int(nrow(exp_dat), 5000),
c(1, 9:18, 20:24, 28:35, 43, 50:61)]/5)

# Get the diagnostic scatter plot before compensation pairs
(exp_dat, pch=".")

# Perform compensation
sce_c <- compCytof(sce, sm, method = "flow", isotope_list=custom_isotope_list)
fs <- sce2fcs(sce_c)

exp_dat <- exprs(fs)
exp_dat <-
asinh(exp_dat[sample.int(nrow(exp_dat), 5000), c(1, 9:18, 20:24, 28:35, 43, 50:61)]/5)
)

# Get the scatter plot after compensation pairs(
exp_dat, pch=".")
```

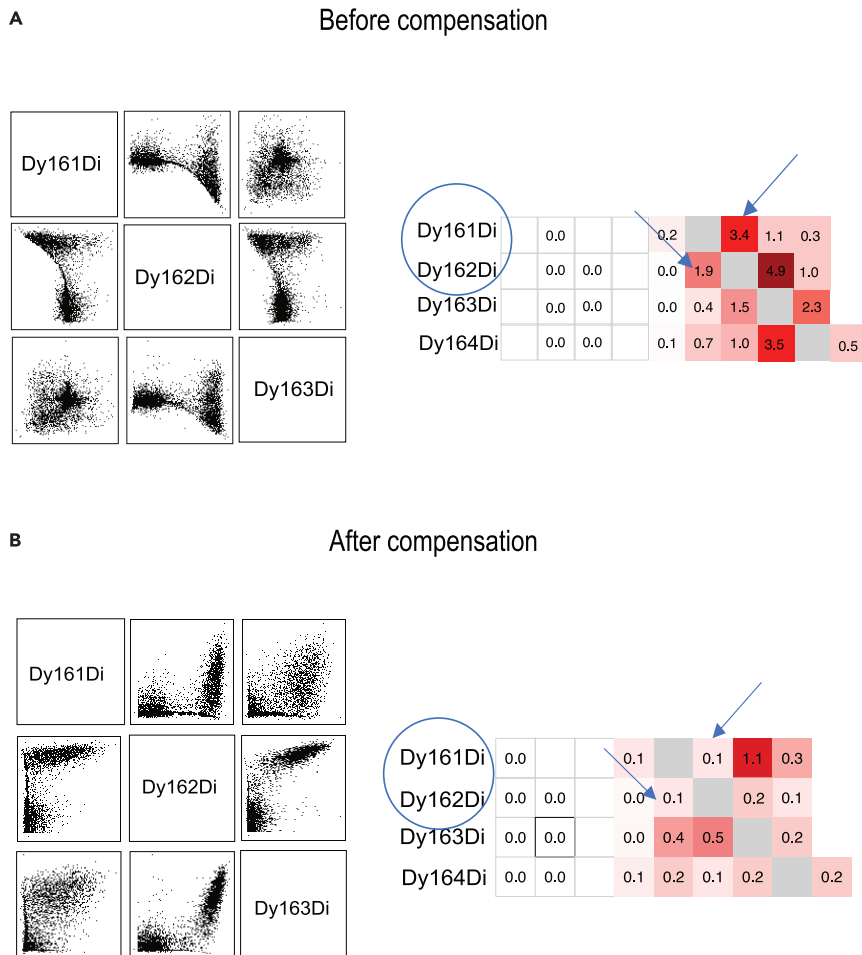


Figure 2. Example of correct compensation of Spillover Matrix and Scatterplots

Have a look at the first Spillover matrix obtained and try to decrease the values for each channel where needed. The rows represent the x-axis, and the columns represent the y-axis. Manually adjust the values for individual channels in the upper triangle by decreasing them.

Note: This step will generate a scatterplot matrix for the signal in all non-compensated channels. It is important to carefully examine the plot to proceed effectively with the final result of compensation.

☐☐ **Pause point:** Take a close look at this example!

The [Figure 2](#) depicts an “over-compensated” channel. The spillover value on the matrix needs to be decreased.

iv. Modify the spillover matrix, compensate, and plot again:

```
# e.g.
# change comp matrix for individual channels
sm[ "Pr141Di" , "Nd142Di" ] <-0.000
sm[ "Nd143Di" , "Nd142Di" ] <-0.001 # new value
sm[ "Nd142Di" , "Nd143Di" ] <-0.001
```

```
sm[ "Nd142Di" , "Gd158Di" ] <-0.002
sm[ "Gd158Di" , "Nd142Di" ] <-0.002
sm[ "Nd142Di" , "Nd144Di" ] <-0.002
sm[ "Nd143Di" , "Nd144Di" ] <-0.001
sm[ "Nd143Di" , "Sm147Di" ] <-0.000
sm[ "Sm147Di" , "Nd143Di" ] <-0.000
sm[ "Nd143Di" , "Tb159Di" ] <-0.002
sm[ "Nd143Di" , "Nd145Di" ] <-0.001
```

- v. Obtain the Spillover matrix with the new values.

```
# New spillmatrix with corrected values
metadata(re)$spillover_matrix <-sm plotSpillmat(re, isotope_list=custom_isotope_list)

# Compensate again with the corrected spillover matrix
sce_c <-compCytof(sce, sm, method = "flow", isotope_list=custom_isotope_list) fs
<-sce2fcs(sce_c)

exp_dat <-exprs(fs)
exp_dat<-

asinh(exp_dat[sample.int(nrow(exp_dat), 5000), c(1, 9:18, 20:24, 28:35, 43, 50:61)]/5)

# Obtain the new scatter plot pairs (exp_dat, pch=".")
```

- vi. Based on the new spillover matrix compensate all FCS files, previously uploaded in the working directory.

```
# files to compensate
files <-dir(pattern=".fcs$")

# you may remove the fcs with single stains
# files <-files[!files %in% single_stains]

# compensate each file with "NNLS" method and save under new for (file in files) {
ff_exp <-flowCore::read.FCS(file, transformation=F, truncate_max_range=FALSE)
ff_exp<prepData(ff_exp)
ff_exp <compCytof(ff_exp, sm, method = "nnls", isotope_list=custom_isotope_list)
ff_exp <sce2fcs(ff_exp)
write.FCS(ff_exp, sub(".fcs", "_comped.fcs", file))
}
```

Data analysis

© Timing: 4+ weeks

In this section we highlight the steps to perform an unsupervised analysis to generate metaclusters using the FlowSOM and ConsensusClusterPlus algorithms, along with the Uniform Manifold Approximation and Projection (UMAP) for dimensionality reduction.

10. Clustering and UMAP visualization.

- a. Load the package dependencies required for the Clustering and UMAP visualization:


```
library(CATALYST)
library(flowCore)
library(flowWorkspace)
library(CytoML)
library(edgeR)
library(flowFP)
library(parallel)
library(diffcyt)
library(scater)
library(ggplot2)
library(RColorBrewer)
library(readxl)
```

b. Load the flowset.

```
# Make sure that the folder "Files" is in the working directory
# Check and set your working directory
getwd()
path <- "Files"
# Get the path to each fcs file
(fcs.files <- dir(path=path, full.names = FALSE))
fs <- read.flowSet(paste0(path, "/", fcs.files), transform=FALSE, truncate_max_range=FALSE)
```

c. Data organization and metadata object:

Create your matrix (called in our script “md”) as an Excel file containing all the features you want to visualize in the analysis ([Table 1](#)).

The “md” must contain the following columns:

file_name, sample_id and “condition”. In longitudinal studies involving a large cohort of patients, the dataset (md) may contain numerous clinical variables. Therefore, carefully select the specific “condition” of interest for the analysis. Clarify your “condition” in this step (e.g., group, body compartment, diagnosis, treatment, gender or age). Save the “md” in the folder that contains the FCS files and the script. Load the “md” in the R environment.

```
# Load metadata stored in an excel file.
md <- read_excel("md.xlsx")
# Select the columns from the md that contained the info and clinical data that the users want to analyze.
md$sample_id <- factor(md$sample_id) # depending on your analysis md$condition_id <- factor(md$condition_id)
md$diagnosis_id <- factor(md$diagnosis_id)
```

d. Look at the description of the parameters stored in the flowset and extract the information. In this section, you will dive into the preliminary information stored in the flowset (the FCS data used). Therefore, carefully examine them, as this step is essential to understand the dataset and will be used in following analysis.

```
# Keep the parameter description fs.desc <- parameters(fs[[1]])@data[,1:2]
# Select channels of interest
```

```

umap.ch.idx<-c(6,13:27,31:38,45,51:56,58:62)

# make marker names more readable and remove unwanted chars p.desc <-unnname(parameters(fs[[1]])$desc)

# Update flowSet with marker names

for (f in 1:length(fs)) { parameters(fs[[f]])$desc < p.desc
}

# Update the parameter description

fs.desc <-cbind(fs.desc, p.desc, umap=logical(nrow(fs.desc))) fs.desc$umap[umap.ch.idx] <- TRUE

fs.desc

fsApply(fs, colnames)

# Get an overview and an estimate of the...
(n.fr <-length(fs)) # ...number of samples

(v.events <-fsApply(fs, nrow)) # ...number of events per sample
(min.events <-min(v.events)) # ...minimum number of events sum(v.events)

cbind(md, v.events)

```

Note: The isotope, metal and antigen information is stored in the flowSet (the container for multiple samples) object.

e. Create the Panel and address the marker information.

In this step, we explore how to create the Panel, which is a data.frame containing columns with the name for each marker present in the input raw data, the targeted protein markers and the marker class ("type" or "state").

Note: It is important to double-check the marker class specification to achieve a robust clustering analysis. FlowSOM/ConsensusClusterPlus will use the "type" markers to perform the clustering. The "state" markers will be considered as functional markers. Markers referred to as "Type" mainly determine phenotypic differences between cell clusters and are typically the lineage markers. The rest of the markers are listed as "State" and are then used to analyze differential marker expression of each cluster between conditions.

f. Depending on your experiment change the name of the marker_class.

```

# Channels and marker names

fcs_colname <-unnname(fs.desc$name)
antigen <-unnname(fs.desc$p.desc, )

# Define the marker classes

# Note: All "type" markers will be used used for clustering marker_class <-rep("none",
nrow(fs.desc))

# Select here the markers needed to be named as "type"
marker_class[c(6,13:27,31:38,45,51:56,58:62)] <- "type"

marker_class <-factor(marker_class, levels = c("type", "state", "none"))

# Create the Panel

panel <-data.frame(fcs_colname, antigen, marker_class, stringsAsFactors = FALSE)

```

Table 1. Example of a possible matrix(md) used to read the flowset

sample_id	condition_id	diagnosis_id	file_name
001	Non-Neuroinflammatory diseases	CON	CON_001.fcs
002	Non-Neuroinflammatory diseases	CON	CON_002.fcs
003	Neuroinflammatory diseases	AD	AD_003.fcs
004	Neuroinflammatory diseases	MS	MS_004.fcs
005	Neuroinflammatory diseases	DEM	DEM_005.fcs

g. Additional information:

```
# Switch the "type" markers as "state" markers if it is needed rowData(sce)$marker_class
[c(seq(10,43),49)] <- "state"

# Double check! rowData(sce)$marker_class
```

h. Create a *SingleCellExperiment* (sce).

This section shows how to store all data used and returned throughout the differential analysis in an object of the *SingleCellExperiment* (sce) class.

Be aware:

The function `prepData()` requires the filenames listed in the `md$file_name` column to match those in the flowSet.

```
# Prepare the Data

md$file_name<-c(keyword(fs, "FILENAME")) # Construct SingleCellExperiment

sce <-prepData(fs, panel, md, features = panel$fcs_colname, md_cols = list(file = "file_name",
id = "sample_id", factors =
c("sample_id","condition_id","diagnosis_id")), panel_cols = list(channel = "fcs_col-
name", antigen = "antigen", class = "marker_class"))
```

△ CRITICAL: You may encounter an error if there is a mismatch between the lists in the "md" and the components in the "md_cols" you are creating. The names of your "condition" in both have to be identical.

i. Visualization of the results with CATALYST package:

Note: This section explores how to obtain results using CATALYST pipeline¹⁹. The details of the procedures are available in <https://github.com/HelenaLC/CATALYST>. Overall, the pipeline allows obtaining a comprehensive explanatory view of the sample dataset through the generation of exploratory plots like Multidimensional scaling plot (MDS) and Non redundancy score plot (NRS). The visualization of FlowSOM heatmaps and UMAP plots after the clustering step provides insights into the distribution of immune cell populations into different meta- clusters based on their similarity.

```
# Perform the clustering

sce <- cluster(sce, features = "type", xdim = 10, ydim = 10, maxK = 20, seed = 1234)

# Visualize the marker expression per cluster with FlowSOM heatmap

plotExprHeatmap(sce, features = "type", by = "cluster_id", k = "meta20", bars=TRUE,perc=TRUE)
```

```
# Dimensionality reduction
# run UMAP on at most 500/1000 cells per sample
sce <- runDR(sce, "UMAP", cells = 1e3, features = "type")
# UMAP plot colored by cluster ID
plotDR(sce, "UMAP", color_by = "meta20")
# Save sce:
saveRDS(sce, "project_sce.rds")
```

▮▮ Pause point: It is highly recommended to save the SingleCellExperiment (sce) as an RDS object after performing the clustering analysis and give it a proper name that allows easy recognition of the project to which the analysis belongs. All the information will be stored in the sce object. This avoids repeating the entire analysis from the beginning, prevent loss information, and enables the analysis to be replicable for other users as well.

△ CRITICAL: Before proceeding with the clustering step, it is essential to verify the number of cells present in each sample. It is recommended practice to include only samples containing a suitable number of cells, typically between 5 to 10 cells per sample (subject to individual discretion) and to exclude samples with lower number of cells. This approach ensures the robustness of subsequent analysis. See troubleshooting section, [problem 5](#).

EXPECTED OUTCOMES

In here, we show the workflow strategy to characterize the immune cell population in CSF samples. Successful completion of the protocol ([Figure 3](#)) should enable the generation of different plots for data visualization using the CATALYST pipeline.¹⁶

Herein, we present results from a small cohort of patients with CON (non-neuroinflammatory disease $n = 3$) and neuroinflammatory disease ($n = 4$) and the corresponding data analysis.

The MDS plot shows similarities between samples in an unsupervised manner. On the other hand, the NRS plot identifies the ability of markers to explain the observed variance in each sample. Differences in cell compositions between CON and neuroinflammatory disease can be seen in the frequency plots ([Figure 4C](#)) as well as in the UMAP plots ([Figure 4D](#)). To further evaluate the phenotypic differences of immune cells between the two conditions, we performed clustering analysis using the FlowSOM and *ConsensusClusterPlus* algorithms. A total of ten clusters were identified. Overall, we could identify the major immune cell subsets including classical, non-classical and intermediate monocytes (C4, C3 and C9 respectively) as well as CD4⁺ T cells (C2) and CD8⁺ T cells (C8 and C7). In most cases CD4⁺ T cells constituted the major population, except in two patients in which the classical monocytes were the most abundant population. Notably, we could identify CD19⁺ B cells present in the CSF of MS patients, but they were virtually absent from the rest, in line with our previous observations.¹

Before proceeding further with the analysis, it is good practice to thoroughly examine the features of the dataset. As explained earlier in this protocol (go to step 9. i), it is important to consider which samples to include in the analysis. For instance, you may have outliers or a small number of cells in certain samples. In our analysis, we set a minimum of 10 cells per sample per cluster to consider for the clustering step. The individual operator must take this point into consideration, depending on the type of the dataset and the experiment they are working on.

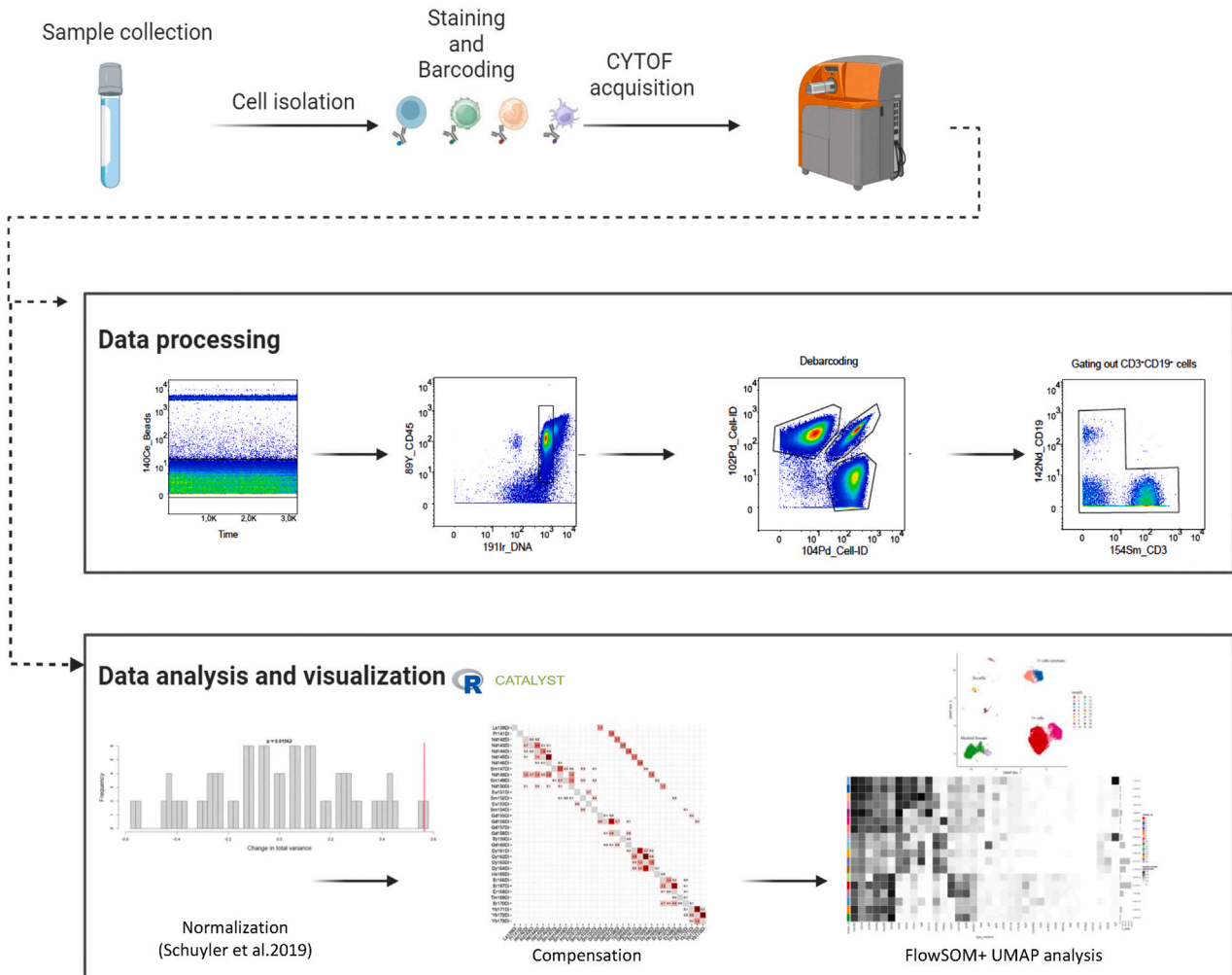


Figure 3. Workflow strategy to characterize immune cell populations by using CyTOF

Figure adapted with [BioRender.com](https://www.biorender.com).

QUANTIFICATION AND STATISTICAL ANALYSIS

The method of our choice is edgeR test which is an optimal statistical test tool for low number of cells. edgeR is a Bioconductor package for differential expression analyses. The package implements exact statistical methods for multigroup experiments developed by Robinson and Smyth. It also implements statistical methods based on generalized linear models (GLMs), suitable for multi-factor experiments of any complexity.

LIMITATIONS

Sample staining and preparation for CyTOF can cause a high rate of cell loss. Typically, only 50–70% of the sample can be recovered in the data. Therefore, studies involving rare cell populations require a larger starting sample size for adequate rigor as compared to studies investigating prevalent subset. The panel design is key to mass cytometry success. However, CyTOF experiments are still practically limited to around 60 markers, meaning that researches must still focus on particular types or functions of cells. Nevertheless, normalization methods and the use of anchor samples can be implemented to improve the reproducibility and comparability of CyTOF results across experiments and study sites.¹ Finally, the CSF studies usually fall short in providing longitudinal data because

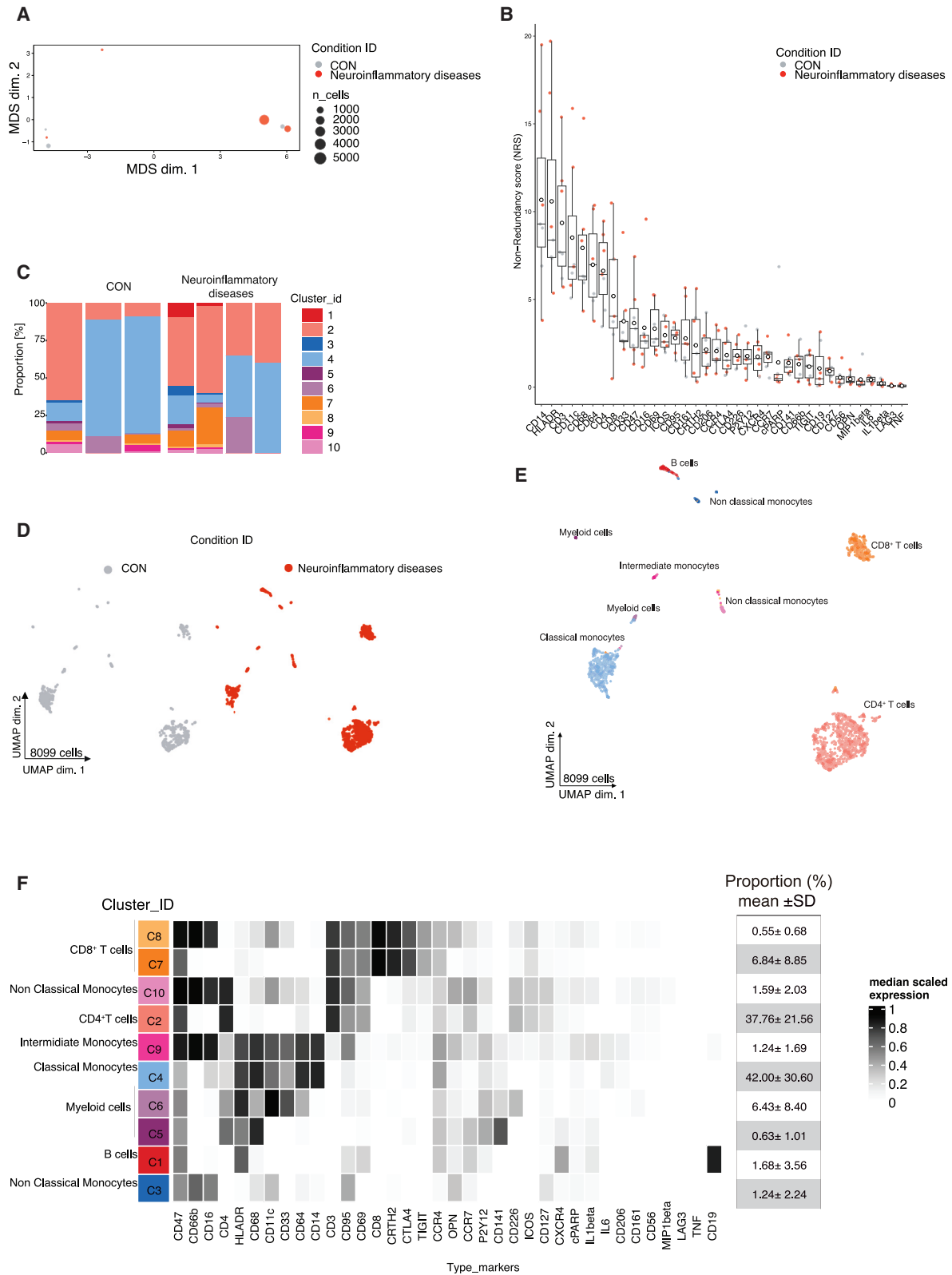


Figure 4. Overview of phenotypical analysis of CSF cells using the described data analysis pipeline

- (A) MDS plot of 7 samples.
- (B) NRS plot. Observed variance of marker expression in each sample. Each dot represents the per-sample NR scores. Whisker plots show the min (smallest) and max (largest) values. The line in the box denotes the median. The empty black circles are mean NR scores.
- (C) Cell population abundance compares the proportions of cell types across the two conditions and aims to highlight populations that are present at different ratios. Bars colored by cluster ID, where the size of a given stripe reflects the proportion of the corresponding cell type in a given sample.
- (D) UMAP projection stratified for condition_id, (E) UMAP projection coloring indicates 1–10 clusters. Each dot represents one cell.
- (F) FlowSOM heatmap depicting the median expression levels of 36 markers for each cluster. Expression levels have been scaled for each marker individually (to the 1st and 5th quintiles) (black, high expression; white, no expression).

repetitive lumbar punctures are difficult to justify. Moreover, defining appropriate control groups is another crucial point because CSF samples from strictly healthy participants are usually not available.¹⁷

However, together with the magnetic resonance imaging measures the CSF cell analysis can improve the diagnostic accuracy and help to estimate individual prognosis.

TROUBLESHOOTING

Problem 1

Due to the small number of cells, the generated pellet could be very small and difficult to see (Step2 before to begin).

Potential solution

We recommended being careful during this passage and always double- checking the supernatant and the pellet. Leave a small amount of the supernatant with the pellet in the Falcon tube.

Problem 2

Contamination of CSF sample (Step2 before to begin). After centrifugation, the sample may contain blood droplets.

Potential solution

Using a CSF blood-contaminated sample will affect the analysis and the reliability of the results. Therefore, we highly recommend excluding any contaminated samples from the analysis.

Problem 3

During the software and library installation, you may encounter a “lib path error” (Step 7). In this case, the dependencies are not recognized by the software.

Potential solution

Make sure to have installed all the software in the same working space and folder on your computer.

Refer to the link below and follow the instructions. If you are downloading the software under a university or institution, ask for the proxy access.

<https://cran.r-project.org/bin/windows/Rtools/rtools40.html>.

```
PATH= "${RTOOLS40_HOME}\usr\bin;${PATH}"

write('PATH= "${RTOOLS40_HOME}\\usr\\bin;${PATH}"', file = "~/Renvirom", append = TRUE)

#Restart R, and verify that make can be found, which should show the path to your Rtools
installation.

Sys.which("make")

## "C:\\rtools40\\usr\\bin\\make.exe"
```

```
#Now try to install the packages on R in order to load them in R studio:  
  
install.packages("jsonlite", type = "source")  
  
#If this succeeds, you're good to go! See the links below to learn more about rtools4 and the Windows build infrastructure.
```

Problem 4

Issue in obtaining the normalized FCS files and the corresponding plots (Step 8.c).

Potential solution

Attempt to run each “channel- marker” individually in a new text file to identify any incorrect channel-marker found in the “in the ChannelToAdjust_example.txt file”.

Problem 5

The number of cells is less than 10 per sample (Step 10.i).

Potential solution

First, check the number of cells in each sample. Finally filter the “sce” object to exclude from the analysis the samples that contain less than 10 cells.

```
# Check the number of cells stored in the sce before running the clustering analysis  
  
# Exclude the samples that don't match the requirement from the dataset sce<-FilterSCE(sce, -  
sample_id != c("filename.fcs"), sample_id !=c("filename.fcs"))  
  
# Perform the clustering  
  
sce <-cluster(sce, features = "type", xdim = 10, ydim = 10, maxK = 20, seed = 1234)
```

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Camila Fernández (maria.fernandez@charite.de).

Technical contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the technical contact, Camila Fernández (maria.fernandez@charite.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

There is no dataset/code associated with the paper.

ACKNOWLEDGMENTS

We would like to thank Christian Böttcher for his technical assistance with sample collection, staining, and mass cytometry experiments. We also thank the BIH Flow & Mass Cytometry Core Facility for their technical assistance and feedback. We would like to thank the TAHRget consortium at ECRC (Experimental and Clinical Research Center -Charité-Universitätsmedizin Berlin, Germany) for supporting the project. Figure was adapted with BioRender.com.

AUTHOR CONTRIBUTIONS

Writing – original draft, G.G.; writing – review and editing, C.F.-Z. and C.B.; methodology, M.W. and C.B.; supervision, project administration, and funding acquisition C.B.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Fernández-Zapata, G., Spruth, M., Gallaccio, D., Dames, J.K., Leman, H., van Dijk, E., Meisel, K., Hol, P., and Priller, B. (2022). Differential compartmentalization of myeloid cell phenotypes and responses towards the CNS in Alzheimer's disease. *Nat. Commun.* 13, 7210. <https://doi.org/10.1038/s41467-022-34719-2>.
- Thygesen, C., Larsen, M.R., and Finsen, B. (2019). Proteomic signatures of neuroinflammation in Alzheimer's disease, multiple sclerosis, and ischemic stroke. *Expert Rev. Proteomics* 16, 601–611. <https://doi.org/10.1080/14789450.2019.1633919>.
- Wong-Guerra, M., Maccioni, R.B., Rojo, L.E., Maccioni, R.B., and Calfio, C. (2023). Revisiting the neuroinflammation hypothesis in Alzheimer's disease: a focus on the druggability of current targets. *Front. Pharmacol.* 14, 1161850. <https://doi.org/10.3389/fphar.2023.1161850>.
- Attfield, K.E., Jensen, L.T., Kaufmann, M., Friese, M.A., and Fugger, L. (2022). The immunology of multiple sclerosis. *Nat. Rev. Immunol.* 22, 734–750. <https://doi.org/10.1038/s41577-022-00718-z>.
- Heming, B., Börsch, A.L., Wiendl, H., and Meyer Zu Hörste, G. (2022). High-dimensional investigation of the cerebrospinal fluid to explore and monitor CNS immune responses. *Genome Med.* 14, 94. <https://doi.org/10.1186/s13073-022-01097-9>.
- Böttcher, S., Sneebouer, K., Knop, P., Fidzinski, K., Snijders, K., Schulz, M., Psy, H., Siegmund, G., Spruth, d.W., Kahn, R.S., et al. (2019). Human microglia regional heterogeneity and phenotypes determined by multiplexed single-cell mass cytometry. *Nat. Neurosci.* 22, 78–90. <https://doi.org/10.1038/s41593-018-0290-2>.
- Böttcher, C., Fernández-Zapata, C., Schlickeiser, S., Kunkel, D., Schulz, A.R., Mei, H.E., Weidinger, C., Gieß, R.M., Asseger, S., Siegmund, B., et al. (2019). Multi-parameter immune profiling of peripheral blood mononuclear cells by multiplexed single-cell mass cytometry in patients with early multiple sclerosis. *Sci. Rep.* 9, 19471. <https://doi.org/10.1038/s41598-019-55852-x>.
- Thrash, E.M., Kleinstaub, K., Hathaway, E.S., Nazzaro, M., Haas, E., Hodi, F.S., and Severgnini, M. (2020). High-Throughput Mass Cytometry Staining for Immunophenotyping Clinical Samples. *STAR Protoc.* 1, 100055. <https://doi.org/10.1016/j.xpro.2020.100055>.
- Rybakowska, P., Burbano, C., Van Gassen, S., Alarcón-Riquelme, M., Aguilar-Quesada, R., Saeys, Y., Alarcón-Riquelme, M.E., and Marañón, C. (2021). Stabilization of Human Whole Blood Samples for Multicenter and Retrospective Immunophenotyping Studies. *Cytometry A.* 99, 524–537. <https://doi.org/10.1002/cyto.a.24241>.
- Fluidigm (2021). *Maxpar Antibody Labeling User guide*.
- Schulz, A.R., Baumgart, S., Schulze, J., Urbicht, M., Grützkau, A., and Mei, H.E. (2019). Stabilizing Antibody Cocktails for Mass Cytometry. *Cytometry A* 95A, 910–916. <https://doi.org/10.1002/cyto.a.23781>.
- Hartwig, A., Arand, M., Epe, B., Guth, S., Jahnke, G., Lampen, A., Martus, H.J., Monien, B., Rietjens, I.M.C.M., Schmitz-Spanke, S., et al. (2020). Mode of action-based risk assessment of genotoxic carcinogens. *Arch. Toxicol.* 94, 1787–1877. <https://doi.org/10.1007/s00204-020-02733-2>.
- McGuire, H.M., and Ashhurst, T.M. (2019). Mass Cytometry, Methods and Protocols (Springer Protocols). <https://doi.org/10.1007/978-1-4939-9454-0>.
- Schuyler, R.P., Jackson, C., Ogolla, S., Rochford, R., Ghosh, D., Rudra, P., Hsieh, E.W.Y., Garcia-Perez, J.E., and Baxter, R.M. (2019). Minimizing Batch Effects in Mass Cytometry Data. *Front. Immunol.* 10, 2367. <https://doi.org/10.3389/fimmu.2019.02367>.
- Chevrier, S., Crowell, H.L., Zanotelli, V.R.T., Engler, S., Robinson, M.D., and Bodenmiller, B. (2018). Compensation of Signal Spillover in Suspension and Imaging Mass Cytometry. *Cell Syst.* 6, 1–9. <https://doi.org/10.1016/j.cels.2018.02.010>.
- Nowicka, M., Krieg, C., Weber, L.M., Hartmann, F.J., Guglietta, S., Becher, B., Levesque, M.P., Robinson, M.D., and Crowell, H.L. (2017). CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. *F1000Res.* 6, 748. <https://doi.org/10.12688/f1000research.11622>.
- Alvermann, H., Stüve, M.D., Wiendl, H., and Stangel, M. (2014). Immunophenotyping of Cerebrospinal Fluid Cells In Multiple Sclerosis. *JAMA Neurol.* 71, 905–912. <https://doi.org/10.1001/jamaneurol.2014.395>.