

Cytometry Part A
Author Checklist: MIFlowCyt-Compliant Items

Requirement	Please Include Requested Information
1.1. Purpose	Test whether or not a linear classifier can automatically and accurately classify different cell types in mass cytometry datasets, and be used to analyze large cohort studies
1.2. Keywords	Single cell, Mass cytometry, Cell type prediction, Machine learning
1.3. Experiment variables	We used five public mass cytometry datasets, including human blood samples, as well as human and mouse bone marrow (described in details in ‘Sample description’)
1.4. Organization name and address	Delft University of Technology Van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands
1.5. Primary contact name and email address	Dr. Ahmed Mahfouz a.mahfouz@lumc.nl
1.6. Date or time period of experiment	From 01-11-2017 to 30-09-2018
1.7. Conclusions	<p>We showed that a linear classifier can be used to automatically assign labels to single cells in mass cytometry data. Using five different datasets, the linear classifier outperforms two state-of-the-art methods (1,2).</p> <p>1. Lee H, Kosoy R, Becker CE, Dudley JT, Kidd BA. Automated cell type discovery and classification through knowledge transfer. <i>Bioinformatics</i> 2017;33:1689–1695. 2. Li H, Shaham U, Stanton KP, Yao Y, Montgomery RR, Kluger Y. Gating mass cytometry data by deep learning. <i>Bioinformatics</i> 2017;33:3423–3430.</p>
1.8. Quality control measures	We used all the published data from the public datasets, and QC is performed in the original studies.
2.1.1.1. (2.1.2.1., 2.1.3.1.) Sample description	<p>We used five publicly available mass cytometry datasets:</p> <ol style="list-style-type: none"> 1. AML is a healthy human bone marrow dataset (3). 2. BMMC is also a healthy human bone marrow dataset (4). 3. PANORAMA dataset consists of 10 replicates of mice bone marrow cells (5). 4. Multi-Center study dataset is a collection of 16 samples drawn from a single subject (2). 5. HMIS is a human mucosal immune system consists of 47 Peripheral Blood Mononuclear Cells (PBMC) samples from individuals with inflammatory bowel diseases (6). Samples are divided into control samples, samples with Crohn’s disease, samples with Celiac disease, and samples with Refractory Celiac disease Type II. <p>2. Li H, Shaham U, Stanton KP, Yao Y, Montgomery RR, Kluger Y. Gating mass cytometry data by deep learning.</p>

	<p>Bioinformatics 2017;33:3423–3430.</p> <p>3. Levine JH, Simonds EF, Bendall SC, Downing JR, Pe D, Nolan GP, Levine JH, Simonds EF, Bendall SC, Davis KL, Amir ED, Tadmor MD, Downing JR, Pe D, Nolan GP. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. Cell 2015;162:184–197. Available at: http://dx.doi.org/10.1016/j.cell.2015.05.047.</p> <p>4. Bendall SC, Simonds EF, Qiu P, Amir ED, Krutzik PO, Finck R, Bruggner R V, Melamed R, Trejo A, Ornatsky OI, Balderas RS, Plevritis SK, Sachs K, Pe’er D, Tanner SD, Nolan GP. Single-Cell Mass Cytometry of Differential Immune and Drug Responses Across a Human Hematopoietic Continuum. Science 2011;332:687–696.</p> <p>5. Samusik N, Good Z, Spitzer MH, Davis KL, Nolan GP. Automated mapping of phenotype space with single-cell data. Nat. Methods 2016;13:493–496.</p> <p>6. van Unen V, Li N, Molendijk I, Temurhan M, Höllt T, van der Meulen-de Jong AE, Verspaget HW, Mearin ML, Mulder CJ, van Bergen J, Lelieveldt BPF, Koning F. Mass Cytometry of the Human Mucosal Immune System Identifies Tissue- and Disease-Associated Immune Subsets. Immunity 2016;44:1227–1239.</p>
2.1.1.2. Biological sample source description	Bone marrow and blood (PBMC) samples
2.1.1.3. Biological sample source organism description	Human and mice samples
2.1.2.2. Environmental sample location	N.A
2.3. Sample treatment description	N.A
2.4. Fluorescence reagent(s) description	N.A
3.1. Instrument manufacturer	N.A
3.2. Instrument model	N.A
3.3. Instrument configuration and settings	N.A
4.1. List-mode data files	We used five public datasets, which are available through the original publications (references above), and through our GitHub link: https://github.com/tabdelaal/CyTOF-Linear-Classifier
4.2. Compensation description	N.A
4.3. Data transformation details	We applied hyperbolic arcsin transformation with a cofactor of 5 for all datasets.
4.4.1. Gate description	N.A
4.4.2. Gate statistics	N.A
4.4.3. Gate boundaries	N.A

Notes

Feel free to use more space than allocated.

You can embed graphics/figures in this document, if needed.

Please make sure to save the document in Microsoft Word version 2003 or older, before uploading to ScholarOne Manuscripts. When uploading this checklist to ScholarOne Manuscripts, please choose the “Supplementary Material for Review” category.

Please note that if your paper is accepted, the checklist will be published as an Online Supporting Information.

For any questions, please contact the Cytometry Part A editorial office at Cytometry@wiley.com.