**Technical Notes**

# PolarFCS: A Multi-Parametric Data Visualisation Aid for Flow Cytometry Assessment

Pavandeep Gill[1], Joanne Luider[1], Etienne Mahe[1 ✉]

[1]University of Calgary, Calgary, Alberta, Canada

Competing interests: PG none; JL none; EM none

## Abstract

Currently available Flow-Cytometry Software (FCS) analysis platforms are computationally efficient and user-friendly, but may lack the functionality of single-plot, multi-parametric data visualisation. Methods to overcome this include gating techniques and/or dimensionality reduction. However, these strategies make Flow-Cytometry (FC) data analysis more time- and labour-intensive; profound errors can also result from incorrect FCS use. We have developed PolarFCS, a software tool capable of single-plot, multi-parametric data visualisation. Unlike traditional clinical FC plots, which typically operate directly on a data-set to produce single-parameter FC histograms or two-parameter orthogonal scatter plots, PolarFCS operates on the flow-parameter calculated centre of mass of each event in the data-set, and presents these as a dot-plot. We compare PolarFCS to our traditional clinical FCS workflow, using a selection of clinical plasma cell FC data. Multiple flow plots and gating strategies are required in the traditional software to isolate neoplastic populations. In PolarFCS, however, positional re-arrangement and scaling of the poles can be used to quickly isolate a population of interest. We also compare both approaches in a case of Minimal Residual Disease (MRD) assessment, and again, the versatility of the polar adjustment and parameter scaling allowable with PolarFCS is demonstrated. PolarFCS employs strategies that allow more accurate, standardised and detailed FC data analysis compared to traditional FCS platforms. Visualisation of multiple parameters in a single plot is an effective and invaluable feature that many other platforms currently do not offer.

Availability: PolarFCS can be downloaded at **https://github.com/etiennemahe/PolarFCS**

## Introduction

Flow cytometry (FC) is a powerful technique through which multiple physical and chemical characteristics (e.g., size, granularity, viability and immunophenotype) of several thousands of cells can be rapidly analysed. In analytical FC, cells in solution are individually passed through an optical flow cell, whereafter scattered incident light can be used to interpolate cellular characteristics. When cells are pre-treated with fluorescently-tagged antibodies, a highly sensitive immunophenotypic signature can also be obtained, exploiting variable incident excitation frequencies and variable emission spectra. Of note, the intensity of light emitted by each antibody-linked fluorophore can be related to antigen expression, and panels containing a number of antibody-linked fluorophores can be easily employed (Verbsky and Routes, 2017). The resulting digitised data are amenable to analysis by FC Software (FCS).

Several FCS platforms are currently available, both as freeware and as commercially available software. While many FCS platforms are computationally efficient and user-friendly, current standard platforms are subject to limitations in functionality. Specifically, the standard means of FCS visualisation involve presentation and analysis of FC data as either single-parameter histograms or two-parameter-correlated plots (either dot plot, density or contour plots) (Ormerod, 2008). As such, these approaches make it impossible to directly visualise data in multi-parameter form (*i.e.*, more than two simultaneous parameters in a single plot). To overcome this, gating techniques are frequently used to analyse subsets of FC data (*e.g.*, using forward scatter and side scatter to analyse only lymphocytes), followed by separate analyses of additional parameters (*e.g.*, lymphocyte CD3 or CD20 intensity) (Ormerod, 2008). These data can then be gated again to produce histograms of even further sub-populations. Gating typically requires the manual isolation of FC-event data in order to highlight cell sub-populations of interest, and requires expert knowledge of the underlying cell populations. Gated FC analysis is a time- and labour-intensive process, and incorrect gating strategies can result in profoundly erroneous FC analysis results, especially when very minor sub-populations are considered. Additionally, a limited two-parameter plot makes it difficult to visualise important multi-parameter relationships that cannot be seen in just two dimensions.

**Technical Notes**

To assist users in overcoming the limitations of traditional FC data analysis, we have developed PolarFCS, a freely-available software tool capable of single-plot, multi-parametric data visualisation. We tested PolarFCS on a selected series of bone marrow specimens submitted for plasma cell FC, and use these exemplars to highlight some of the features that a multi-parametric FC analysis might afford the user.

## Materials, Methodologies and Techniques

### Technical Summary: PolarFCS

PolarFCS provides a means of single-plot, multi-parametric FC analysis. Unlike traditional FC plots, which operate directly on a data-set, producing either single-parameter FC data histograms or two-parameter orthogonal scatter plots, PolarFCS operates on the flow-parameter calculated centre of mass of each event in the data-set. To do this, PolarFCS begins by considering each FC parameter as a polar axis in two-dimensional space. For each event, the signal intensity of each parameter is plotted along the length of each parameter's corresponding polar axis. The result, when these coordinates are subtended by lines, is a polygon in two-dimensional space. A simple algorithm[1] is then applied to determine the polygon's centre of mass. The centres of mass for each event are then presented as a dot-plot.

Akin to other FC data-analysis platforms, PolarFCS also allows users to interact with the data-set to permit more detailed data analysis. PolarFCS is able to incorporate previously determined gating strategies by way of dot-plot colour variation; it also allows users to adjust the orientation of the polar axes. Such adjustments will have the effect of re-positioning certain data subsets relative to others. Such adjustments can also be used to amplify or dampen the contribution of a given parameter relative to the others. Finally, an adjustable counting field allows users to accurately determine the number of events within a given area relative to the total event count. This latter tool aims to assist users in comparing event fractions between studies, or to compare the relative abundance of a given subset of events.

PolarFCS is written in MATLAB[2] and consists of two primary functions. The `masthead` PolarFCS function offers users a basic graphical-user-interface to select data files for input (either as tabular '.txt' or '.csv' data, or as FCS-formatted '.lmd' List Mode Data files[3]), and to select and rename FC data parameters as desired. The second, `makefcspolarscatter`, function can be invoked directly by MATLAB users, requiring a user-defined uncompressed data-set, parameter list, plot title and colour matrix. The colour matrix allows users to encode a

[1]http://www.seas.upenn.edu/~sys502/extra_materials/Polygon%20Area%20and%20Centroid.pdf

[2]https://www.mathworks.com/products/matlab.html

[3]More detailed formatting descriptions are provided at https://github.com/etiennemahe/PolarFCS

specific pre-defined gating strategy in the form of colour variation in the final plot. The PolarFCS Windows- and Mac OSX-compiled stand-alone versions, as well as the source code for the `makefcspolarscatter` function, are available for download at https://github.com/etiennemahe/PolarFCS. Other compilation formats and additional source code can be made available by specific request to the corresponding author.

## Results

### Application Example: PolarFCS as a visual aid in plasma cell FC

The current haematopathology work-up of plasma cell neoplasia relies heavily on FC. Plasma cell FC provides
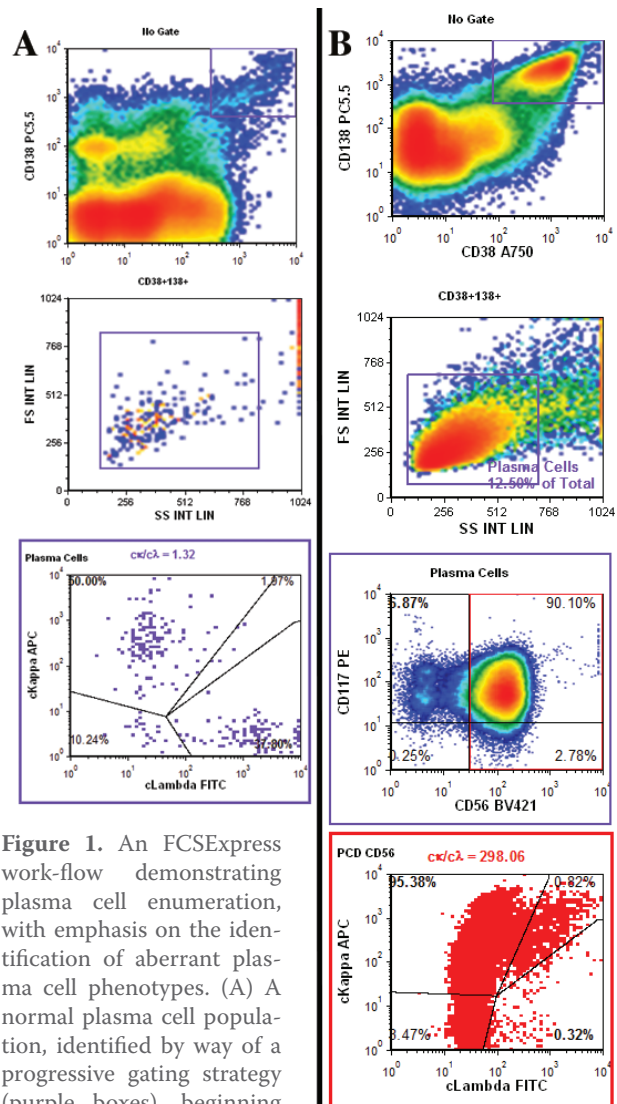


**Figure 1.** An FCSExpress work-flow demonstrating plasma cell enumeration, with emphasis on the identification of aberrant plasma cell phenotypes. (A) A normal plasma cell population, identified by way of a progressive gating strategy (purple boxes), beginning from CD38/CD138-bright events (which are notably few in number), followed by gating around a region of typical plasma cell forward and side-scatter, followed finally by assessment of cytoplasmic kappa/lambda ratios (which are within the normal limits in this case); (B) A neoplastic plasma cell population, identified by way of a comparable progressive gating strategy, with aberrant dual CD56/CD117 expression (red-highlighted gate), and an associated extreme increase in the cytoplasmic kappa/lambda ratio.

Technical Notes
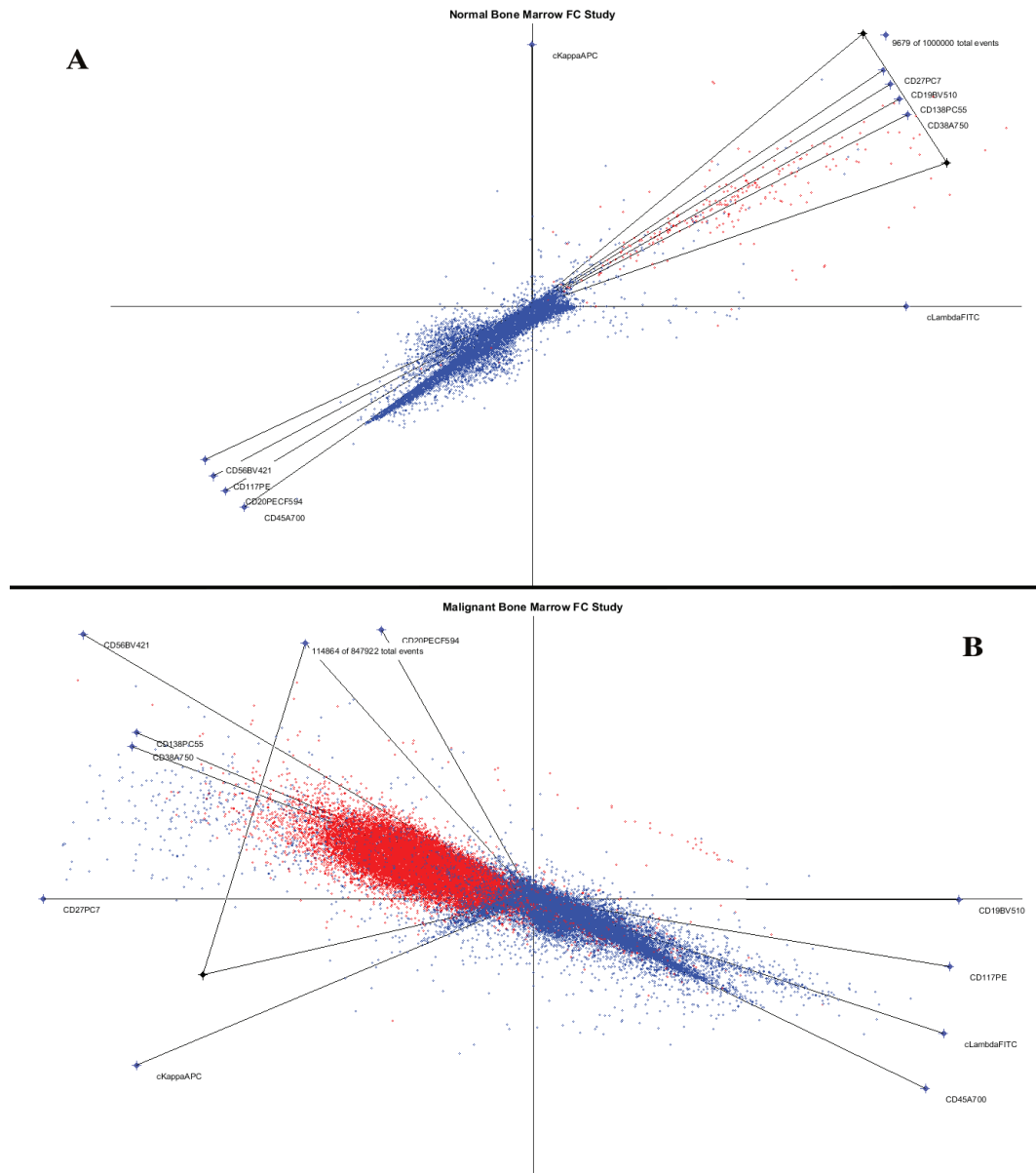


**Figure 2.** PolarFCS demonstrating plasma cell enumeration, with the plasma cells identified by the gating strategies in figure 1 highlighted red. (A) Normal plasma cell population; (B) Neoplastic plasma cell population. Note that positional re-arrangement and scaling of the PolarFCS poles can be used to isolate populations of interest.

a rapid and sensitive quantitation of the plasma cell fraction from bone marrow specimens, and allows rapid and reproducible clonal assessment. FC can be used to identify aberrant immunophenotypes suggestive of neoplasia, and plasma cell FC permits highly sensitive Minimal Residual Disease (MRD) assessment.

To explore how PolarFCS might be applied in these contexts, we identified example clinical cases to compare our standard FC informatic workflow (FCSExpress 5, IVD, DeNovo Software, Glendale, CA[4]) with one employing PolarFCS[5]. For demonstration purposes, we exported the gated populations of interest, as well as the overall non-gated study data, as raw text-formatted data, and
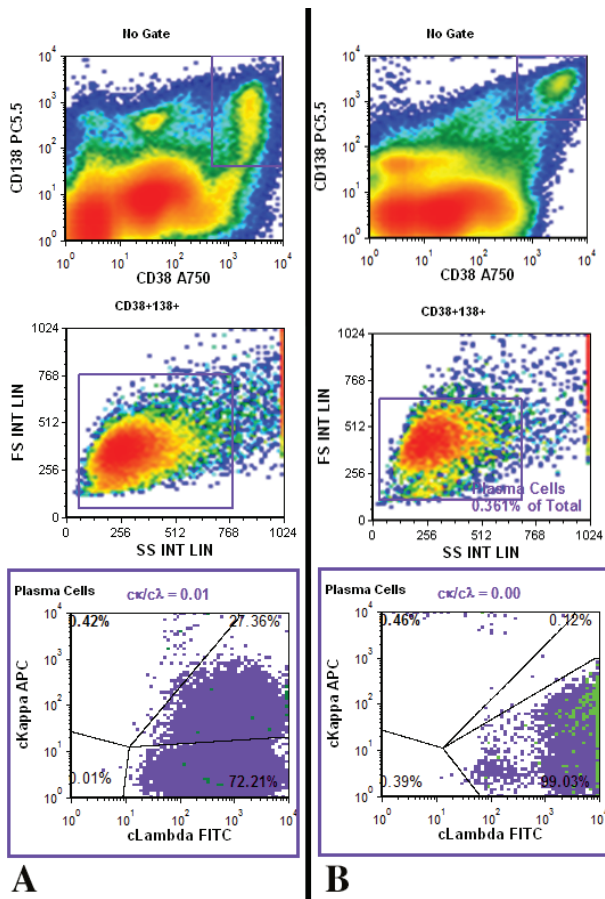
subsequently employed MATLAB's set algebra functions to define a custom colour profile for input into the `makefcspolarscatter` function. In our jurisdiction, the type of exploratory work outlined in this manuscript qualifies as a 'quality assurance or evaluation project' (ARECCI Ethics Screening Tool - 2005, revised 2010[6]), and therefore does not require direct institutional ethics board review. Nevertheless, the tenets of the Declaration of Helsinki, as further required by Canadian Tri-Council Research Ethics Panel (Canadian Institutes of Health Research, 2014[7]), were adhered to rigorously.

[4]https://www.denovosoftware.com/

[5]Some of the de-identified data-sets used in our analyses are available for download at https://github.com/etiennemahe/PolarFCS

[6]http://www.aihealthsolutions.ca/media/Sept-2013-Paper-Version-ScreeningTool.pdf

[7]http://www.pre.ethics.gc.ca/pdf/eng/tcps2-2014/TCPS_2_FINAL_Web.pdf

**Technical Notes**



**Figure 3.** FCSExpress workflow highlight a minimal residual disease work-up. (A) Plasma cell myeloma at diagnosis; (B) Follow-up status post autologous bone marrow transplant.

Plasma cell enumeration, with emphasis on the identification of aberrant plasma cell phenotypes, is shown in Figures 1-2. Figure 1A delineates the orthogonal flow plots required to gate out a small phenotypically normal plasma cell population, whereas Figure 1B highlights an abundant neoplastic (and immunophenotypically aberrant) plasma cell population. In contrast, the appertaining PolarFCS plots are shown in Figure 2 ('Normal' case in Figure 2A and 'Plasma Cell Myeloma' in Figure 2B). This contrast serves to highlight the clear difference between 'normal' and 'malignant' PolarFCS signatures, as well as the versatility of pole adjustment and parameter scaling allowable in PolarFCS. Of note, PolarFCS allows poles corresponding to expected 'normal' plasma cell phenotypic features to be subtended in opposition to 'abnormal' phenotypes; in doing so, PolarFCS allows clear separation of these distinct sub-populations, without an otherwise required multi-step gating strategy.

MRD assessment is highlighted in Figures 3-4. Figure 3 delineates the orthogonal flow plots of a case of plasma cell myeloma, at diagnosis (Figure 3A) and at follow-up status, post autologous bone marrow transplant (Figure 3B). In contrast, the appertaining PolarFCS plots are shown in Figure 4 (at diagnosis in Figure 4A, and at follow-up in Figure 4B). This contrast serves to highlight the persistent disease population identifiable in both

studies, facilitated by positioning of the PolarFCS poles in comparable orientations. The PolarFCS plot in Figure 4B also demonstrates a characteristic 'wall-artefact', partly highlighted in green in the lowest flow plot of Figure 3B. Arguably, this artefact is more readily apparent in the PolarFCS plot. When the original and corresponding PolarFCS gated-event counts are compared, it becomes apparent that the reference flow plot likely over-estimates the gated-event count.
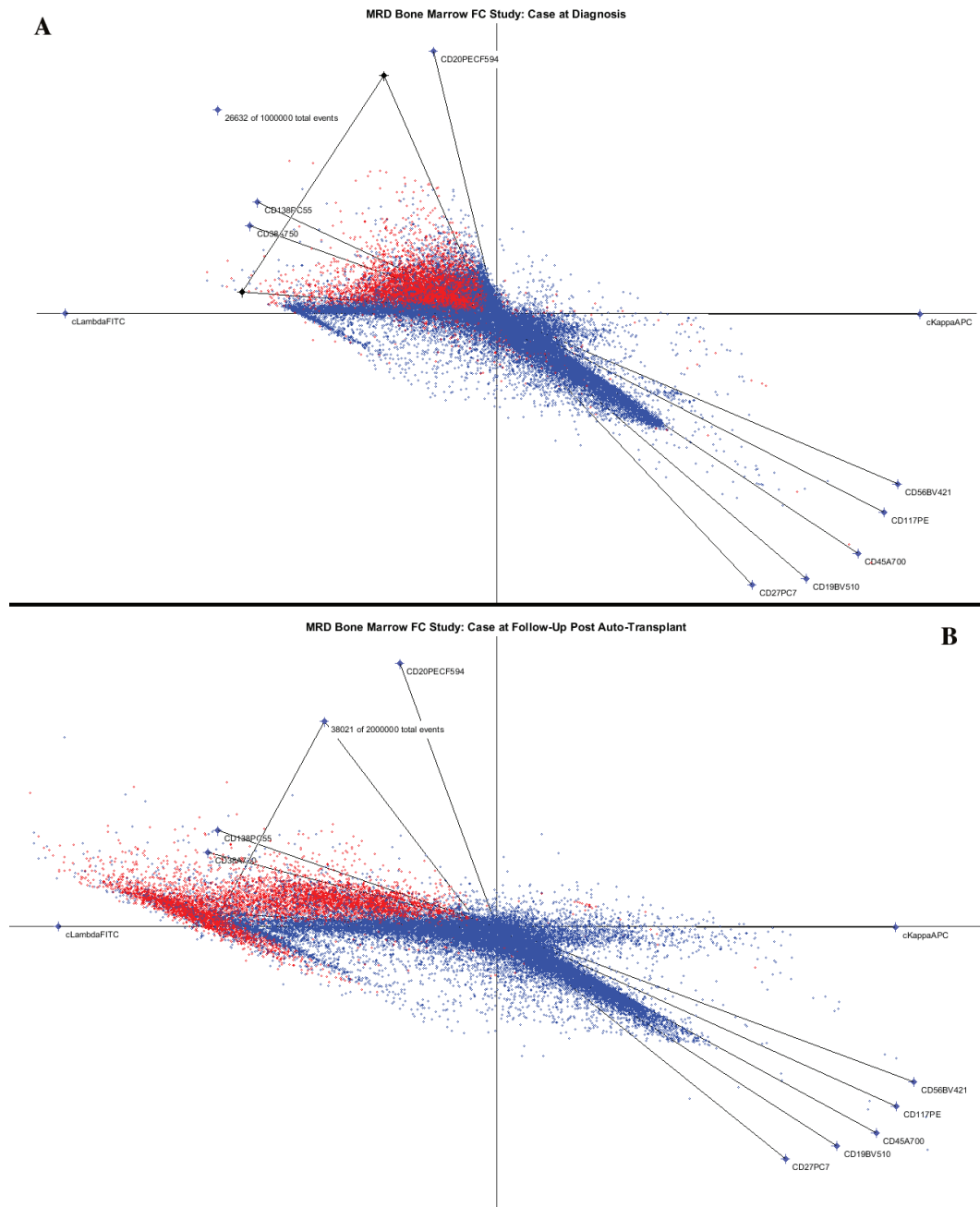
## Discussion

The application of FC to the work-up of haematological malignancies is invaluable. Each year, our laboratory receives several thousand specimens for FC analysis; and, from among these, several hundred cancer cases are diagnosed (Canadian Cancer Society, 2016[8]). Thus, in order to optimise laboratory processes, an efficient FC workflow requires both sensitive and efficient FCS. Current clinical FCS systems might be amenable to additional process improvements, including dimensionality reduction techniques, for visualisation of multiple parameters in a single plot.

Examples of currently available tools for dimensionality reduction include Spanning-tree Progression Analysis of Density-normalised Events (SPADE), Principle Component Analysis (PCA), and viSNE. SPADE works by clustering cells into sub-populations, measuring the average of each cluster, and then presenting the data in a branched-tree structure showing the connections between the different clusters (Qiu *et al.*, 2011). PCA tries to maintain single-event data through linear transformations of relationships between putatively correlated variables; however, this limits PCA in its application to non-linear relationships (Amir *et al.*, 2013). viSNE is a dimensionality-reduction software using the t-distributed Stochastic Neighbour Embedding (t-SNE) algorithm, and has the capacity for single-event data, and visualisation of non-linear relationships (Amir *et al.*, 2013).

With these (and other) dimensionality-reduction tools, there are limitations to how much of the high-dimensional data can be represented in a low-dimensional map. Detail relating to less dominant parameters can be lost using these approaches. In addition, there may be limited user control over which parameters are included or excluded. In contrast, PolarFCS allows users to select the inclusion of parameters, as well as their scaling, by way of polar adjustment. The number of events that can be analysed using viSNE also has an upper limit; as such, viSNE may need to be run on a well-defined, pre-filtered subset of data (possibly requiring additional pre-analytical analysis algorithms). No such data limits are inherent to PolarFCS.

It bears noting that PolarFCS may also be subject to limitations in its current version. For example,

[8]http://www.cancer.ca/~/media/cancer.ca/CW/cancer%20information/cancer%20101/Canadian%20cancer%20statistics/Canadian-Cancer-Statistics-2016-EN.pdf?la=en

**Figure 4.** PolarFCS applied to the cases highlighted in Figure 3. (A) Plasma cell myeloma at diagnosis; (B) Follow-up status post autologous bone marrow transplant. Note that PolarFCS polar axes can be arranged in comparable orientations between experiments, to permit visual (and numerical) comparison of FC changes over time. Also highlighted in this case is the very obvious "wall artefact" which has likely unknowingly skewed the FCSExpress results, but which can be easily highlighted in PolarFCS.

the current PolarFCS software lacks the capacity to incorporate successive gating strategies, a limitation also found with viSNE. However, as has been emphasised, such gating strategies can be time-consuming and error-prone, and may not be necessary in multi-parametric analyses. Also, PolarFCS has not yet been tested across the broader array of clinical scenarios mandating FC analysis. That said, in the limited settings in which we have employed it, as outlined herein, PolarFCS appears to provide a visually satisfying adjunct to current FCS solutions.

In conclusion, we proffer PolarFCS as a novel and simple FC data-analysis system that offers single-plot, multi-parametric FC data visualisation and analysis. PolarFCS has been compiled to be compatible with both Windows and MacOSX operating systems, which we hope will allow wide availability.

**EMBnet**.journal
Bioinformatics in Action

Technical Notes

### Key Points

- PolarFCS is an FCS program that allows easy single-plot, multi-parametric data visualisation, a feature that traditional FC data-analysis programs currently do not offer.
- We compare PolarFCS to a traditional FCS platform in the analysis of haematological malignancies, and show how the visualisation functionalities available in PolarFCS allow more accurate, standardised and detailed data analysis.
- PolarFCS has highly valuable characteristics, especially as most hospital laboratories receive thousands of specimens for FC each year, and traditional FC analysis is a time- and labour-intensive process.

## Acknowledgements

## References

1. Amir ED, Davis KL, Tadmor MD, Simonds EF, Levine JH *et al.* (2013) viSNE enables visualization of high dimensional single cell data and reveals phenotypic heterogeneity of leukemia. *Nat. Biotechnol.* **31**(6), 545-552. http://dx.doi.org/10.1038/nbt.2594
2. Ormerod MG (2008) Flow Cytometry – A Basic Introduction. 1st Ed. De Novo Software. http://flowbook.denovosoftware.com (accessed on January 31, 2017).
3. Qiu P, Simonds EF, Bendall SC, Gibbs KD, Bruggner RV *et al.* (2011) Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nat. Biotechnol.* **29**, 886-891. http://dx.doi.org/10.1038/nbt.1991
4. Verbsky J and Routes JM (2017) Flow cytometry for the diagnosis of primary immunodeficiencies. In: UpToDate, Notarangelo LD (Ed), UpToDate, Waltham, MA. Updated February 2017: http://www.uptodate.com/contents/flow-cytometry-for-the-diagnosis-of-primary-immunodeficiencies (accessed on February 28, 2017).