

# Implementing flowDensity for Automated Analysis of Bone Marrow Lymphocyte Population

Ghazaleh Eskandari, Sishir Subedi, Paul Christensen, Randall J. Olsen, Youli Zu, Scott W. Long

Department of Pathology and Genomic Medicine, Houston Methodist Hospital, Houston, TX, USA

Submitted: 02-02-2021

Accepted: 27-09-2021

Published: 09-12-2021

## Abstract

**Introduction:** Manual gating of flow cytometry (FCM) data for marrow cell analysis is a standard approach in current practice, although it is time- and labor-consuming. Recent advances in cytometry technology have led to significant efforts in developing partially or fully automated analysis methods. Although multiple supervised and unsupervised FCM data analysis algorithms have been developed, they have not been widely adopted by the clinical and research laboratories. In this study, we evaluated flowDensity, an open source freely available algorithm, as an automated analysis tool for classification of lymphocyte subsets in the bone marrow biopsy specimens. **Materials and Methods:** FlowDensity-based gating was applied to 102 normal bone marrow samples and compared with the manual analysis. Independent expression of each cell marker was assessed for comprehensive expression analysis and visualization. **Results:** Our findings showed a correlation between the manual and flowDensity-based gating in the lymphocyte subsets. However, flowDensity-based gating in the populations with a small number of cells in each cluster showed a low degree of correlation. Comprehensive expression analysis successfully identified and visualized the lymphocyte subsets. **Discussion:** Our study found that although flowDensity might be a promising method for FCM data analysis, more optimization is required before implementing this algorithm into day-to-day workflow.

**Keywords:** Automated analysis, data analysis, data visualization, flow cytometry, flowDensity, gating

## INTRODUCTION

Flow cytometry (FCM) has been extensively used to identify clusters of cells that share a particular expression pattern of surface and intracellular proteins.<sup>[1-4]</sup> Manual gating, a subjective process of sequential inspection of one or two characteristics at a time, is currently the gold standard for FCM data analysis.<sup>[3,5]</sup> This manual process suffers from a lack of standardization and is found to be a significant source of variation in the FCM studies, with interlaboratory coefficients of variations up to 30%.<sup>[5]</sup> Other limitations associated with manual gating include subjectivity, difficulties in detecting unknown cell populations, and difficulties in reproducibility. To standardize this process and address these crucial limitations, major efforts have been made to develop partially or fully automated analysis methods.<sup>[1]</sup> Automated analysis methods can be grouped into two

main categories: unsupervised and supervised.<sup>[1]</sup> Most of the automated gating methods are unsupervised. In working with the unsupervised algorithms, the data do not come with predefined labels and there are no outputs to predict. In terms of FCM data, these algorithms find some characteristic features of cells and they use these features to group the cells into different populations. As a result of this approach, previously unknown cell populations can also be identified in an unbiased, data-driven manner. The unsupervised algorithms do not require any training and need no or very limited parameterization, which makes

**Address for correspondence:** Dr. Scott Wesley Long, Associate Professor of Pathology, Department of Pathology and Genomic Medicine, Houston Methodist Hospital, 6565 Fannin St, M227, Houston, TX 77030, USA.  
E-mail: SWLong@houstonmethodist.org

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

**For reprints contact:** reprints@medknow.com

**How to cite this article:** Ghazaleh Eskandari, Sishir Subedi, Paul Christensen, Randall J. Olsen, Youli Zu, Scott W. Long. Implementing flowDensity for automated analysis of bone marrow lymphocyte population. J Pathol Inform 0;0:0.

### Access this article online

#### Quick Response Code:



**Website:**  
www.jpathinformatics.org

**DOI:**  
10.4103/JOPI.JOPI\_12\_21

them easy to use.<sup>[1,5,6]</sup> In contrast, supervised approaches start with the goal of predicting a known output. They use external variables such as external biological or clinical characteristics as an input to train a model, which can then be used to predict the status of the new samples.<sup>[1,6]</sup>

Despite the new breakthroughs in the FCM bioinformatics and development of several algorithms for automated analyses of FCM data, these methods have not been widely adopted by the clinical and research laboratories. This limited adoption could be due to the lack of bioinformatics expertise required for implementing these tools, failure of these algorithms in replicating a human expert's gating results, or difficulties in selecting the most appropriate method.<sup>[3,7,8]</sup> Multiple studies comparing various automated FCM data analysis techniques have been published. In one of the series of publications by FlowCAP ("Flow Cytometry: Critical Assessment of Population Identification Methods") Consortium, seven different combinations of FCM data analysis algorithms were evaluated and two out of seven approaches showed promising outcomes.<sup>[1,7]</sup> flowDensity is one of the two successful methods used in this study, and it is a freely available, open-source, supervised clustering algorithm that closely matches an expert's sequential two-dimensional (2D) gating strategy to identify the predefined cell populations.<sup>[3]</sup> The sequence of gates is specified by the user and customized threshold calculations based on one-dimensional (1D) density estimation are used for different cell subsets.<sup>[3,4]</sup> The algorithm is developed by using R, a free statistical software platform widely used for biological analyses. Using 1D thresholds makes flowDensity computationally efficient, and parallel processing of several files can be performed on a standard desktop computer.<sup>[4]</sup>

The development of automated FCM data analysis methods has made using multiple data analysis and visualization techniques possible. Some of these methods are also incorporated into the currently available software for manual analysis and interpretation of FCM data. One of the common visualization methods for high-dimension data, such as in FCM, is principal component analysis (PCA). PCA is a technique for reducing the dimensionality of large data sets into a biaxial plot that retains most of the variation in the data and aids in identifying different cell clusters based on the gene expression.<sup>[9,10]</sup> In the present study, the flowDensity algorithm was implemented as an automated analysis tool for the evaluation of lymphocyte subsets in the bone marrow biopsy specimens. PCA was then applied for visualization of the cell populations identified by flowDensity. In addition, a comprehensive expression analysis pipeline was developed for cytometric profiling and investigation.

## MATERIALS AND METHODS

### Data collection

Overall, 102 bone marrow biopsy specimens collected from January 1, 2019 to December 31, 2019 were included

in this retrospective study, after ethics approval from the Institutional Review Board (IRB) at the Houston Methodist Hospital. The biopsies had been performed as part of a clinical workup to rule out various bone marrow abnormalities and were all diagnosed as "normal for age" by a board-certified hematopathologist. The basic bone marrow lymphocyte panel in our institution includes the following surface markers: CD2, CD3, CD4, CD5, CD7, CD8, CD10, CD19, CD20, CD34, CD38, CD45, CD56, and Kappa and Lambda light chains. The FCM data were acquired by using FACSDiva software (BD Biosciences, San Jose, CA) and stored as FCS files after proper compensation. All the cases had been previously reviewed by a hematopathologist using manual gating and the following cell populations had been identified: CD4+ T cells, CD8+ T cells, Natural killer (NK) cells (CD3-/CD56+/low SSC), NK-like T large granular lymphocytes (NK-like T LGL) (CD3+/CD56+/low SSC), CD56+ cells (with low and high SSC), CD19+/Kappa+ B cells, CD19+/Lambda+ B cells, CD10+ cells (with low SSC), CD19+/Kappa+/CD10+ cells, CD19+/Lambda+/ CD10+ cells, CD19+/CD5+ B cells, plasma cells (CD38+), and CD34+ cells.

### Data processing with flowDensity

The FCS files were read and transformed by using R programming language (flowCore package).<sup>[11]</sup> The log transformation (logbase = 10) was applied on all the markers, and cells with positive values were selected for further analysis. An iterative training model was used to optimize the flowDensity gating parameters for each marker. First, about one-third of the total samples (35 out of 102 samples) were randomly selected and tested by using the default parameters. The results of the automated gating were then visually inspected by a pathologist and compared with the manual gating done by a hematopathologist. An acceptable threshold was defined as achieving comparable results between the automated and manual gating in at least 30 out of 35 cases. Therefore, this step was repeated until the threshold was met by using a different set of parameters during each iteration. These parameters are algorithm features used by flowDensity, such as density distribution-shape, percentile, location, etc. Once the gating strategy was configured, a single R script was prepared to gate all the markers in each tube. No further modification was applied to the finalized script to improve the outcomes for individual samples. The output files contained density plot, gated plot, and a CSV file with a total number, proportion, and identifier of cells selected for each combination of markers used for gating in each tube (e.g., CD3+/low SSC).

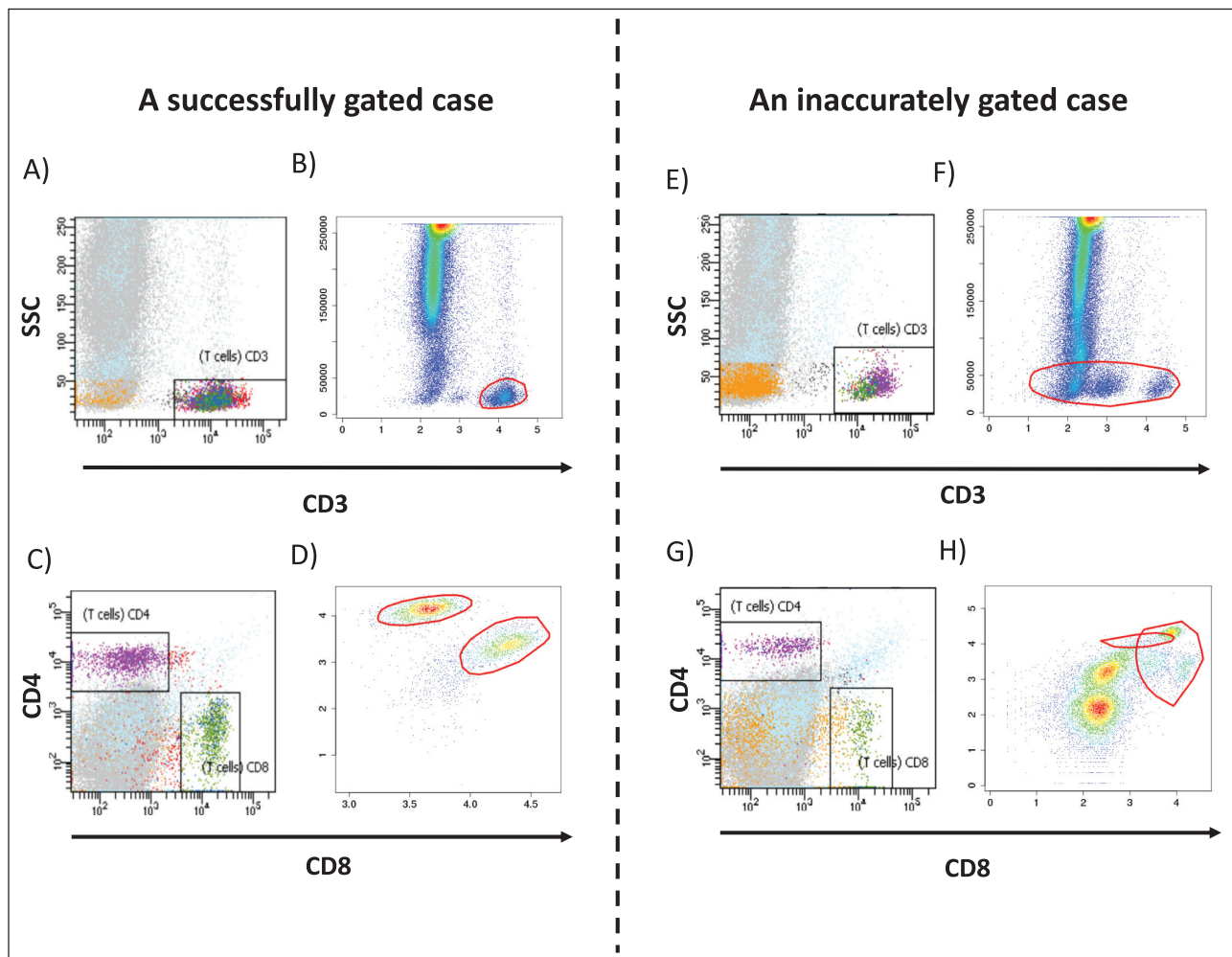
### Data analysis and visualization

The flowDensity results with an optimized gating strategy from 102 specimens were rigorously evaluated by two board-certified pathologists and categorized into pass and fail groups based on the accuracy of the distribution

estimation and gating [Figure 1]. In addition, the cell proportions for each gated cluster from the cases that passed the initial assessment were correlated with its corresponding manual proportions.

The cytometric profile of each cell in a tube consists of expression values for multiple markers. Thus, each cell can be projected onto an  $n$ -dimensional space, where each dimension is represented by a marker. To visualize such data in multidimensional space, we used PCA, a dimensionality-reduction method that linearly combines a large number of parameters (expression values from different markers) and derives a smaller set of pseudoparameters while preserving most of the information present in the original high-dimensional data. Here, we used PCA to visualize selected cell populations expressing different markers identified by the optimized automated workflow described earlier.

Traditionally automated FCM tools such as flowDensity are used to identify specific subpopulation of cells using sequential parent gating. These tools can be also utilized to develop an automated workflow that can identify all the subset of cells in a tube expressing a different set of markers to a single cell resolution. Here, we expanded the flowDensity algorithm to conduct a comprehensive expression analysis to profile all the subpopulations of cells expressing a unique set of markers in a tube. To achieve this, we first reoptimized the gating strategy independently for each marker as described in the method section but without any sequential parent gating. Then, we generated a data matrix for each tube where each row is identified by a cell and each column is identified by a combination of two markers used for gating the cell populations (e.g., CD3+/low SSC). The matrix  $X$  has an entry  $X_{ij}$  as 1 if cell  $i$  is gated by the algorithm as positive for the marker pair  $j$ , else  $X_{ij}$  is 0. This binary matrix was used as Boolean gating



**Figure 1:** Manual and flowDensity-based gatings on two representative samples. (A–D) An example of successfully gated case (classified as Pass). The gates are placed manually (Plot A) and by using flowDensity (Plot B) on the CD3+/low SSC cell population. This population was then used as the parent gate and subsequently gated as CD4+ and CD8+ manually (Plot C) and by flowDensity (Plot D). (E–H) An example of inaccurately gated case (classified as fail). The gates are placed manually (Plot E) and by using flowDensity (Plot F) on the CD3+/low SSC cell population. This population was then used as the parent gate and subsequently gated as CD4+ and CD8+, manually (Plot G) and by flowDensity (Plot H)

to generate all possible subsets of cells expressing a unique combination of markers in each tube.

The downstream data analysis and visualization were conducted by using the Python programming language (sklearn package used for PCA).<sup>[12]</sup> The pipeline code is publicly available at <https://github.com/Houton-Methodist-Clinical-Informatics/flow-cytometry-data-analysis>.

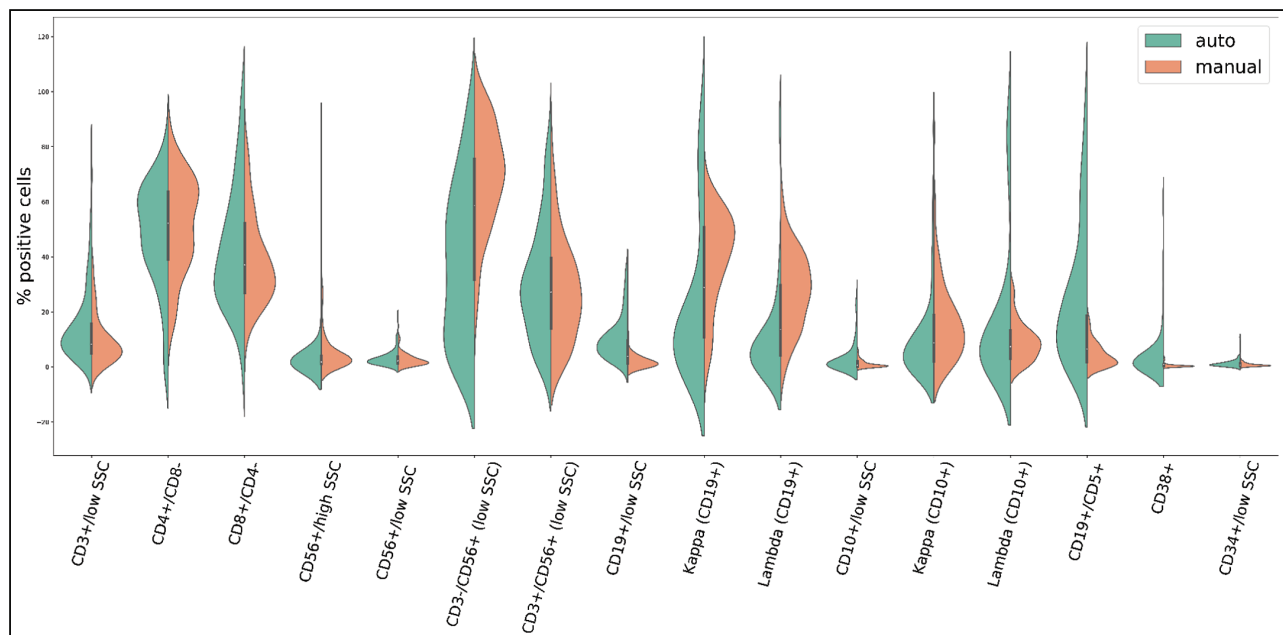
## RESULTS

### Gating on populations of interest

Automated gating was performed for 102 normal bone marrow specimens, on the same populations as the manual gating. In comparing flowDensity and manual gating results, the proportion distribution of cells from all 102 cases showed an overlap within a 95% confidence interval between the two methods for all the markers; however, for 6 out of 16 markers (CD38+, Lambda [CD10+], CD10+/low SSC, CD3-/CD56+ [low SSC], CD19+/low SSC, and CD19+/CD5+), the equality of variances in the distributions was significantly different (Levene's test,  $P_{val} < 0.01$ , Figure 2).

As seen in the Method section, all the automatic gating scattograms were then reviewed by two board-certified pathologists and classified as pass or fail based on the accuracy of gating. For example, cases with a too large or narrow selection boundary or inappropriate boundary shape were classified as fail [Figures 1 and 3]. There were

no discrepancies in evaluating the accuracy of gating between the two pathologists. Out of 102 normal bone marrow specimens, 95 cases were successfully gated for the CD3+/low SSC cell population and showed a strong correlation ( $r = 0.9843$ ) [Figure 4A]. Sequential gating on CD3-positive cells for CD4 and CD8 showed 90 and 71 correctly gated cases, respectively ( $r = 0.9865$  and  $0.9509$ ) [Figure 4B and C]. Overall, 97 out of 102 specimens with successful gating on CD56+/low SSC ( $r = 0.9001$ ) were subsequently processed for NK (CD3-/CD56+/low SSC) and NK-like T LGL (CD3+/CD56+/low SSC) populations. In 97 specimens gated for NK, 68 samples were successful ( $r = 0.6478$ ) [Figure 4D]. The number of correctly processed cases was higher in gating the NK-like TLGL cells (92 out of 97) and showed a Pearson correlation coefficient of 0.8792 [Figure 4E]. Gating on the CD56+/high SSC cell population resulted in a lower number of correctly gated samples (52 out of 102,  $r = 0.6934$ ). Using flowDensity to gate the B cell subsets showed a lower number of successfully gated cases (56 out of 102 for CD19+/low SSC) with a Pearson correlation coefficient of 0.5455 [Figure 4F]. The successfully processed cases were subsequently gated for Kappa and Lambda and showed 30 and 36 correctly gated samples ( $r = 0.5111$  and  $0.5329$ , respectively) [Figure 4G and H]. CD19+/low SSC cells were also used as the parent gate for gating CD5+ cells and showed 37 cases correctly gated ( $r = -0.0012$ ). In gating the CD10+/low SSC population, 84 out of 102 specimens were successfully processed ( $r = 0.8712$ ) [Figure 4I]. Sequential gating for Kappa and Lambda was



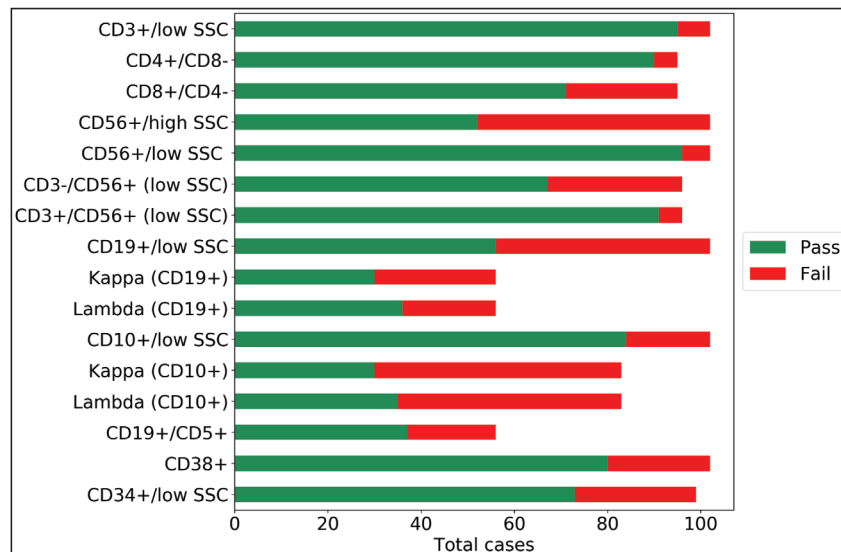
**Figure 2:** Comparing manual and automated-based gatings. Violin plots showing the distribution of cell proportions using automated (flowDensity) and manual gatings. The proportion distribution of cells from all 102 cases showed an overlap within a 95% confidence interval between the two methods for all the markers; however, for 6 out of 16 markers (CD38+, Lambda (CD10+), CD10+/low SSC, CD3-/CD56+ (low SSC), CD19+/low SSC, and CD19+/CD5+), the equality of variances in the distributions was significantly different (Levene's test,  $P_{val} < 0.01$ )

performed and showed 30 and 35 successfully processed samples, respectively ( $r = 0.6729$  and  $0.2988$ ). Automated gating on CD38+ cells on 102 samples showed 80 correctly gated cases ( $r = 0.2018$ ). Gating CD34-positive cells had a Pearson correlation coefficient of 0.4945, with 73 out of 99 being successfully processed [Figure 4J].

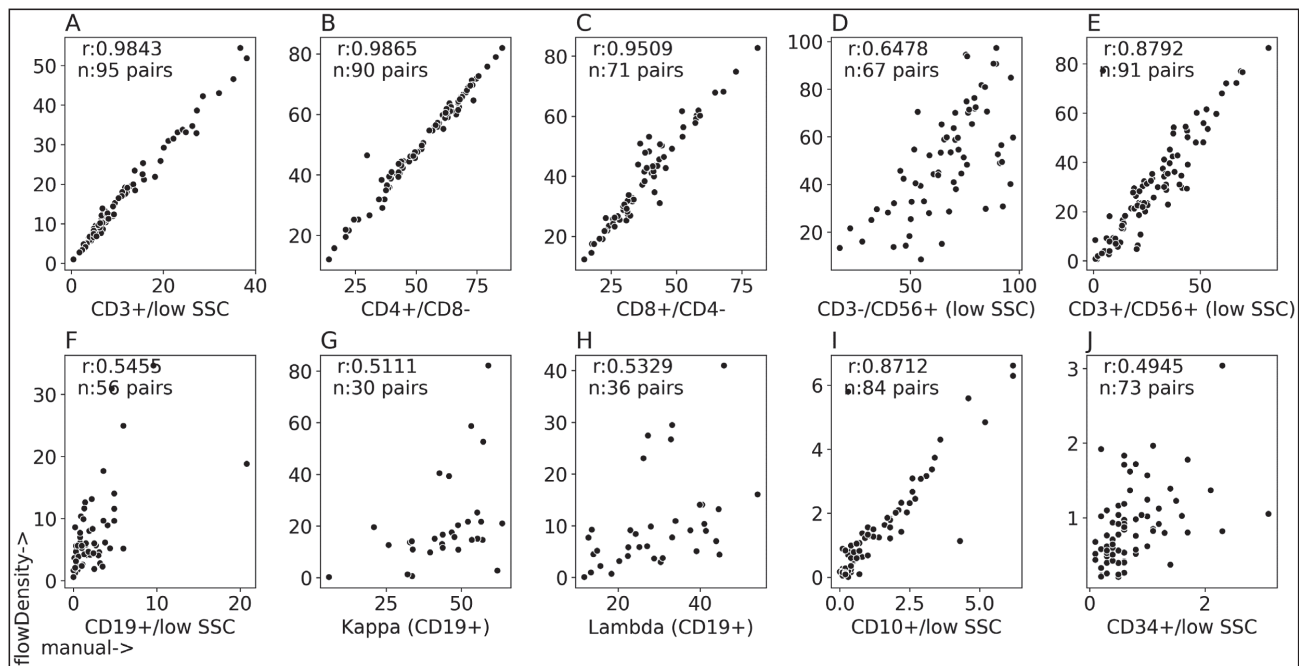
### Applying principal component analysis

To visualize the cell clusters, the PCA technique was applied to 50 samples that were successfully gated by flowDensity

for the main cell subsets in the T cell (CD4+ T cells, CD8+ T cells, NK cells, and NK-like TLGL) or CD19+ B cells (Kappa+ B cell, Lambda+ B cell, and CD5+ B cells) screening tubes. The PCA algorithm assigned each cell in a new dimension (such as PC1 and PC2), where each new dimension is a linear combination of the expression values of all the markers present in a tube. These two dimensions (PCA1/PCA2) are plotted on a biaxial plot to visualize clusters of cells and compared against subsets of cells identified by flowDensity for the respective cell types. Here,



**Figure 3:** Total number of cases gated by flowDensity. All the gated plots were reviewed by a pathologist, and the samples that were successfully processed were identified. This chart shows the total number of passed and failed cases for each cell population individually



**Figure 4:** Correlation of flowDensity and manual gatings. Comparing the manual and flowDensity-based gatings. Correlation plots for the following representative cell populations are shown: CD3+ (A), CD4+ (B), CD8+ (C), CD3-/CD56+ (D), CD3+/CD56+ (E), CD19+ (F), Kappa+ (G), Lambda+ (H), CD10+ (I), CD34+ (J)

the FCM data from 28 T cell and 22 B cell tube samples, successfully analyzed by flowDensity, were used for PCA visualization and evaluation. By applying PCA in our study, the cells of the same subset identified by flowDensity analysis were mostly grouped and separated from the cells of different subsets, although CD5+ B cell subsets with a few number of events were not readily distinguishable from the adjacent clusters. Comparing the visualization results of the T cell subsets showed a consistent pattern in 20 out of 28 normal bone marrow samples. However, this ratio was lower in the B cell subpopulations (9 out of 22). The PCA-generated maps for six representative T and B cell samples are shown in Figure 5. Figure 5A shows the PCA-generated map visualizing T cell subsets in three different bone marrow specimens; Figure 5B depicts B cell subpopulations of three representative bone marrow samples.

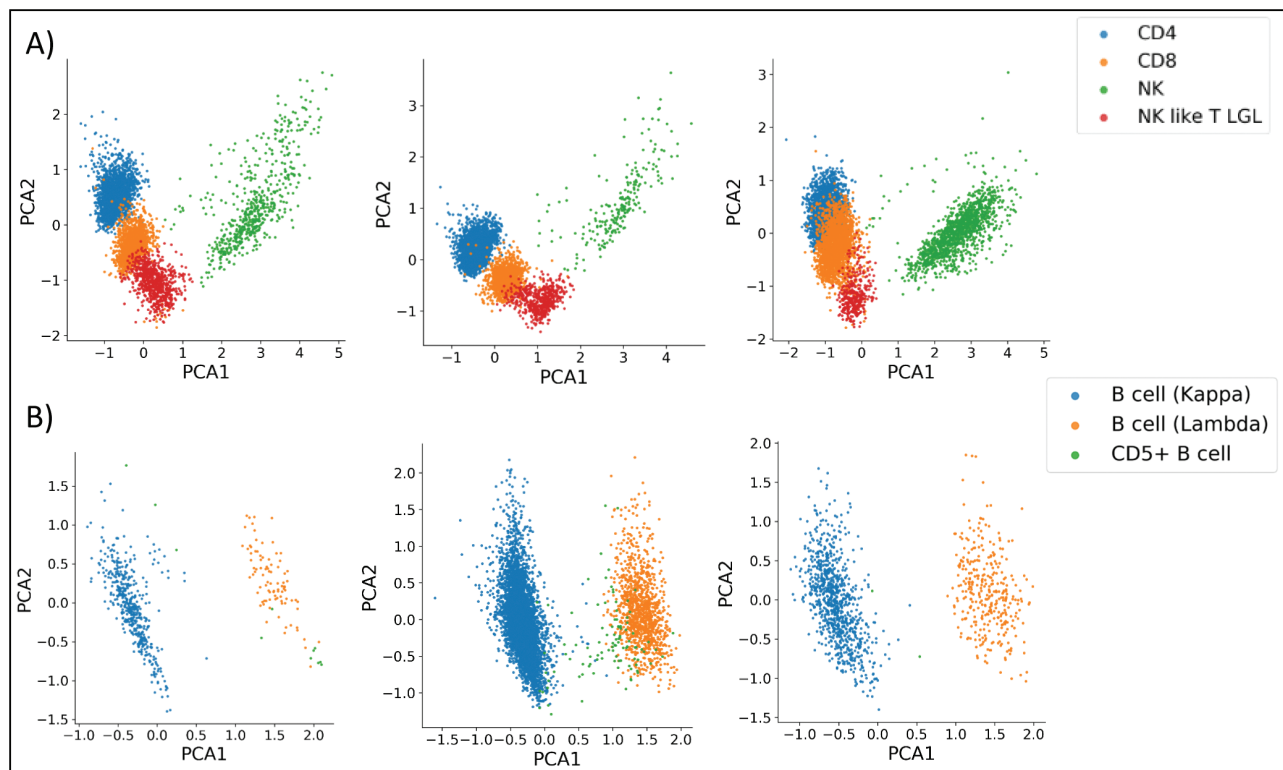
### Independent analysis of individual cells and 2D plot visualization

To identify all the subpopulations present in each sample, we utilized the same gating strategy with flow density as described earlier but removed all the parent gatings from the pipeline. With this approach, we were able to analyze each cell independently and identify a set of markers that it expressed. This information was then used to identify all the

subpopulations present in the entire sample and calculate the total number of cells for each subset. Next, each cell was assigned to its group marked by a set of markers that it expressed, and the cell clusters with greater than 0.1% of total events were then visualized on 2D scatterplots [Figure 6].

### DISCUSSION

Recent advancements in cytometry technology with an increased number of measured parameters per cell and increased data complexity have made automated analysis tools a potential solution for handling the abundance of produced data.<sup>[7]</sup> flowDensity, a supervised algorithm for analyzing FCM data, has recently been used in different studies.<sup>[3,4,13]</sup> Conrad *et al.*<sup>[4]</sup> used flowDensity to implement an automated FCM analysis pipeline for human immune profiling. Data were generated by performing two staining panels for the identification of effector and memory or helper and regulatory T cells, and they showed a strong correlation between the manual and automated methods.<sup>[4]</sup> In another study by Ivison *et al.*,<sup>[13]</sup> an automated analysis workflow using flowDensity was developed and peripheral blood samples from both healthy subjects and patients 10 days after hematopoietic stem cell transplantation were analyzed. Data were acquired by using different instruments from different vendors and across centers. The results were

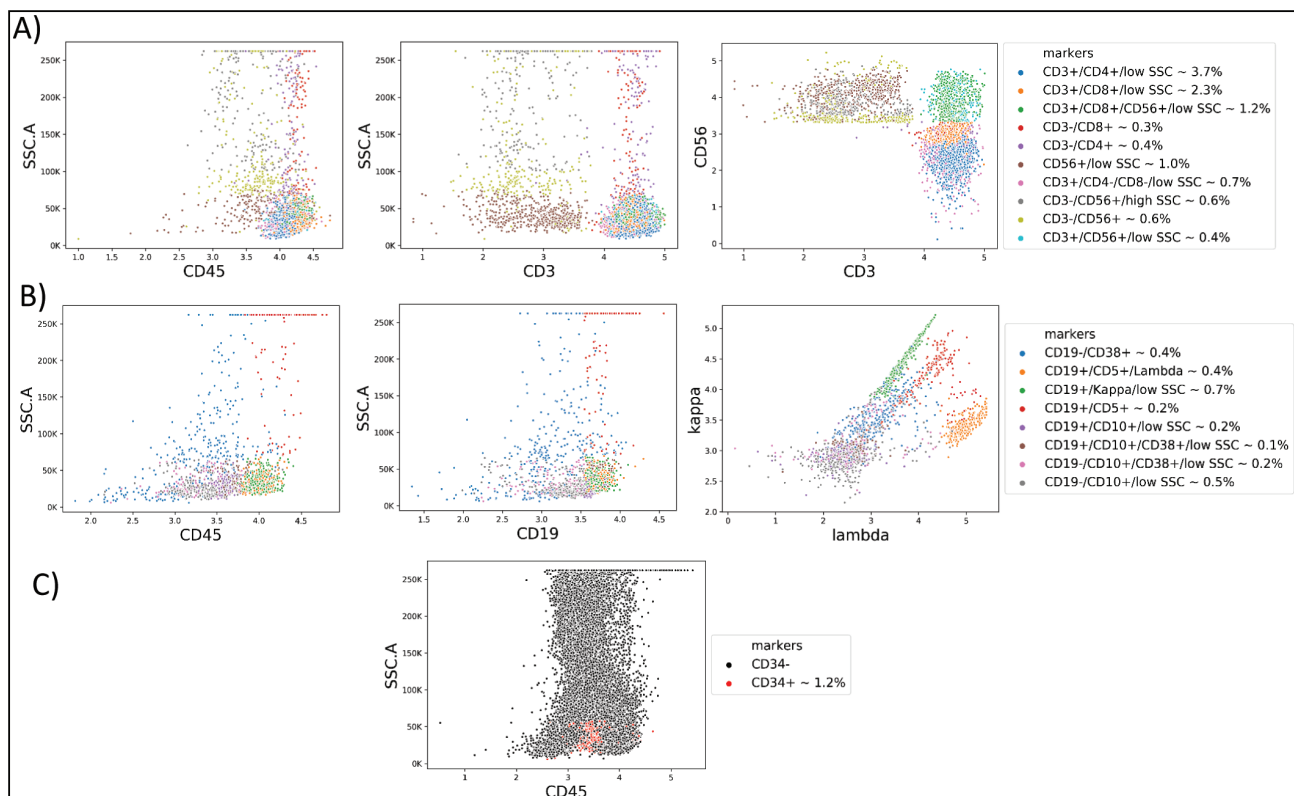


**Figure 5:** PCA-generated maps. By applying the PCA technique to the bone marrow FCM data and visualizing the subpopulations (CD4+ T cells, CD8+ T cells, NK cells, and NK-like TLGL, Kappa+ B cell, Lambda+ B cell, CD5+ B cells), the cells of the same subpopulation identified by flowDensity analysis were successfully grouped and separated from the other cells. (A) Shows the PCA generated map visualizing T cell subsets in three different bone marrow specimens. The B cell subpopulations of three representative bone marrow samples are shown in (B)

compared with per-event value obtained by an expert manual analyzer and showed a strong agreement between the automated and manual methods.<sup>[13]</sup> In our study, we implemented flowDensity algorithm to analyze the FCM data of bone marrow biopsy lymphocyte subsets. For the populations that were successfully gated by flowDensity (pass group), our results showed a strong correlation between the manual and automated methods for the majority of the lymphocyte subsets. However, the flowDensity-based gating method failed to create an accurate gate for 5%–50% of the cell populations [Figure 3]. In addition, automated gating of the cell populations with a small number of events did not match the manual gating as closely. This finding was in agreement with the previous publications identifying the clusters with a few events to show a larger discrepancy between the manual and flowDensity-based gateings.<sup>[4,13]</sup> Other contributors of disagreement between the two methods in recent studies include poorly defined clusters and populations with indistinct marker expression boundaries, which are also known to be a significant source of variability in manual gating.<sup>[4,5,13]</sup> By applying PCA, the main cell subsets (CD4+ T cells, CD8+ T cells, NK cells and NK-like TLGL, Kappa+ B cell, and Lambda+ B cell) identified by flowDensity were successfully grouped and separated from the cells of different subsets. In addition, T cell subsets showed a consistent pattern in 20 out of 28

normal bone marrow samples. However, this ratio was lower (9 out of 22) in visualizing the B cell subpopulations. This finding could be due to the presence of a smaller number of B cells in the bone marrow, which likely reduces the reliability of a flowDensity-based approach. Further, to identify all the present cell populations in each sample, we defined the expression of the variant markers for each cell independently. All the present cell populations with a unique combination of markers were identified, and the clusters with greater than 0.1% of total events were visualized on 2D scatterplots [Figure 6]. This approach helped us to readily identify all the present cell populations, calculate their proportion in the sample, and easily recognize them on the created plots. We believe that using this approach in the FCM workflow can facilitate the process of FCM data analysis and interpretation in both research and clinical practice.

In summary, major efforts have been made in recent years for the development of automated FCM data analysis, in order to increase efficiency and standardization of data output and reduce the analysis time. Although several reports of successful implementation of automated gating methods on either artificial data sets or patients' samples have recently been published, there is still limited adoption due to barriers such as the required bioinformatics expertise.<sup>[4,5,14-16]</sup> Our study found that although



**Figure 6:** Independent analysis of individual cells in a representative case. By analyzing each cell independently, we were able to identify all the present cell populations with a unique combination of markers. All the cell subsets with greater than 0.1% of total events were visualized on 2D scatterplots. (A) Three representative plots showing the cell populations present in the T cell screening tube. (B) Three representative plots visualizing the cell populations present in the B cell screening tube. (C) CD34+ cells from the blast tube

flowDensity might be a promising method for FCM data analysis, more optimization and potentially incorporating additional tools for identification of clusters with a small number of cells are still required before implementing this algorithm in our everyday practice. Increasing the number of studies using clinical FCM data will provide additional feedback to the algorithm developers and improve accuracy in the future versions. In addition, the development of user-friendly tools that require less informatics and programming knowledge can result in a more widespread adoption of these tools.

### Authors' contributions

Ghazaleh Eskandari: Conception or design of the work, data collection, data analysis and interpretation, drafting the article, and final approval of the version to be published. Sishir Subedi: Conception or design of the work, data collection, data analysis and interpretation, drafting the article, and final approval of the version to be published. Paul Christensen: Conception or design of the work, critical revision of the article, and final approval of the version to be published. Randall J. Olsen, MD: Conception or design of the work, critical revision of the article, and final approval of the version to be published. Youli Zu, MD, PhD: Conception or design of the work, critical revision of the article, and final approval of the version to be published. S. Wesley Long, MD, PhD: Conception or design of the work, critical revision of the article, and final approval of the version to be published.

### Financial support and sponsorship

Nil.

### Conflicts of interest

There are no conflicts of interest.

### REFERENCES

1. Weber LM, Robinson MD. Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data. *Cytometry Part A* 2016;89:1084-96.
2. Menon V, Thomas R, Ghale AR, Reinhard C, Pruszk J. Flow cytometry protocols for surface and intracellular antigen analyses of neural cell types. *J Vis Exp* 2014:e52241.
3. Malek M, Taghiyar MJ, Chong L, Finak G, Gottardo R, Brinkman RR. Flowdensity: Reproducing manual gating of flow cytometry data by automated density-based cell population identification. *Bioinformatics* 2015;31:606-7.
4. Conrad VK, Dubay CJ, Malek M, Brinkman RR, Koguchi Y, Redmond WL. Implementation and validation of an automated flow cytometry analysis pipeline for human immune profiling. *Cytometry A* 2019;95:183-91.
5. Brinkman RR. Improving the rigor and reproducibility of flow cytometry-based clinical research and trials through automated data analysis. *Cytometry A* 2020;97:107-12.
6. Deo RC. Machine learning in medicine. *Circulation* 2015;132:1920-30.
7. Aghaeepour N, Chattopadhyay P, Chikina M, Dhaene T, Van Gassen S, Kursu M, *et al.* A benchmark for evaluation of algorithms for identification of cellular correlates of clinical outcomes. *Cytometry A* 2016;89:16-21.
8. Aghaeepour N, Finak G, Hoos H, Mosmann TR, Brinkman R, Gottardo R, *et al.* FlowCAP Consortium; DREAM Consortium. Critical assessment of automated flow cytometry data analysis techniques. *Nat Methods* 2013;10:228-38.
9. Jolliffe IT, Cadima J. Principal component analysis: A review and recent developments. *Philos Trans A Math Phys Eng Sci* 2016;374:20150202.
10. Ringnér M. What is principal component analysis? *Nat Biotechnol* 2008;26:303-4.
11. Ellis B, Haaland P, Hahne F, Le Meur N, Gopalakrishnan N, Spidlen J, *et al.* flowCore: FlowCore: Basic structures for flow cytometry data. R package version 2.4.0. 2021. Available from: <http://bioconductor.org/packages/release/bioc/html/flowCore.html>. [Last accessed on 2021 Oct 20].
12. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B. Scikit-learn: Machine learning in Python. *J Mach Learn Res* 2011;12:2825-30.
13. Ivison S, Malek M, Garcia RV, Broady R, Halpin A, Richaud M, *et al.* A standardized immune phenotyping and automated data analysis platform for multicenter biomarker studies. *JCI Insight* 2018;3:e121867.
14. Burel JG, Qian Y, Lindestam Arlehamn C, Weiskopf D, Zapardiel-Gonzalo J, Taplitz R, *et al.* An integrated workflow to assess technical and biological variability of cell population frequencies in human peripheral blood by flow cytometry. *J Immunol* 2017;198:1748-58.
15. Chen X, Hasan M, Libri V, Urrutia A, Beitz B, Rouilly V, *et al.*; Milieu Intérieur Consortium. Automated flow cytometric analysis across large numbers of samples and cell types. *Clin Immunol* 2015;157:249-60.
16. Lacombe F, Lechevalier N, Vial JP, Béné, MC. An R-derived flowSOM process to analyze unsupervised clustering of normal and malignant human bone marrow classical flow cytometry data. *Cytometry Part A* 2019;95:1191-7.