Cytometry Part B (Clinical Cytometry) 80B:282–290 (2011)

Original Article

Immunophenotypic Signatures of Benign and Dysplastic Granulopoiesis by Cytomic Profiling

William G. Finn,[1]* Alexandra M. Harrington,[2] Kevin M. Carter,[3] Raviv Raich,[4] Steven H. Kroft,[2] and Alfred O. Hero III[5]

[1]Department of Pathology, University of Michigan, Ann Arbor, Michigan
[2]Department of Pathology, Medical College of Wisconsin, Milwaukee, Wisconsin
[3]Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts
[4]Department of Electrical Engineering and Computer Science, Oregon State University, Corvallis, Oregon
[5]Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan

Background: The role of flow cytometry (FCM) in diagnosing myelodysplastic syndromes (MDS) remains controversial, because analysis of myeloid maturation may involve subjective interpretation of sometimes subtle patterns on multiparameter FCM.

Methods: Using six-parameter marker combinations known to be useful in evaluating the myeloid compartment in MDS, we measured objective immunophenotypic differences between non-neoplastic (n = 25) and dysplastic (n = 17) granulopoiesis using a novel method, called Fisher information nonparametric embedding (FINE), that measures information distances among FCM datasets modeled as individual high-dimensional probability density functions, rather than as sets of two-dimensional histograms. Information-preserving component analysis (IPCA) was used to create information-optimized "rotated" two-dimensional histograms for visualizing myelopoietic immunophenotypes for each individual sample.

Results: There was a consistent trend of segregation of higher-grade MDS (RAEB and RCMD) from benign by FINE analysis. This difference was accentuated in cases with morphologic dysgranulopoiesis and in cases with clonal cytogenetic abnormalities. However, lower grades of MDS or cases that lacked morphologic dysgranulopoiesis showed much greater overlap with non-neoplastic cases. Two cases of reactive left shift were consistently embedded within the higher-grade MDS group. IPCA yielded two-dimensional histogram projections for each individual case by relative weighting of measured cellular characteristics, optimized for preserving information distances derived through FINE.

Conclusions: Objective analysis by information geometry supports the conclusions of previous studies that there are immunophenotypic differences in the maturation patterns of benign granulopoiesis and high grade MDS, but also reinforces the known pitfalls of overlap between low-grade MDS and benign granulopoiesis and overlap between reactive granulocytic left shifts and dysplastic granulopoiesis.

Key terms: myelodysplastic syndromes; flow cytometry; cytomics; machine learning

How to cite this article: Finn WG, Harrington AM, Carter KM, Raich R, Kroft SH, Hero AO. Immunophenotypic signatures of benign and dysplastic granulopoiesis by cytomic profiling. Cytometry Part B 2011; 80B: 282–290.

The role of flow cytometric immunophenotyping in the diagnosis of myelodysplastic syndromes (MDS) is controversial. There is an emerging consensus that myelodysplastic hematopoiesis displays immunophenotypic aberrancies that may be helpful in diagnosis (1–4). However, the numerous studies on this topic have often used different approaches, and there is no single standard for immunophenotypic assessment of myelodysplasia. Furthermore, most studies rely on individual

Grant sponsor: National Science Foundation; Grant number: CCR-0325571.
*Correspondence to: William G. Finn, University of Michigan, Department of Pathology, Room M5242 Medical Science I, 1301 Catherine Road, Ann Arbor, MI 48109-5602, USA.
E-mail: wgfinn@umich.edu
Received 21 September 2010; Revision 25 January 2011; Accepted 10 February 2011
Published online 15 March 2011 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/cyto.b.20592

interpretation of patterns displayed on two-dimensional projections of higher dimensional flow cytometry (FCM) data, often using subjective descriptions of subtle shape differences between dysplastic and benign hematopoiesis on such plots (3–6). Although in many ways the subjective interpretation of such patterns (as opposed to assessment of individual marker expression by often arbitrary quantitative thresholds) is a step forward-analogous to the subjective interpretation of histopathologic patterns in surgical pathology, this approach also raises concern over interobserver reproducibility in the interpretation of myelodysplastic versus benign hematopoiesis and renders such interpretations susceptible to observer bias.

Recently, we described a method of objectively measuring differences in the high-dimensional data distributions created by multicolor FCM datasets (7). This method, known as Fisher information nonparametric embedding (FINE), uses the principles of information geometry to determine information distances between individual high-dimensional FCM list mode files, with each file treated as a single high-dimensional distribution or object, rather than as a series of two-dimensional dotplots (7,8). A related method, information preserving component analysis (IPCA), allows direct visualization of an information-optimized two-dimensional histogram for each case, derived from all measured FCM markers based on the differences in information among different cases (9).

In this study, we use the FINE and IPCA algorithms to compare the immunophenotypic signatures of benign and myelodysplastic myelopoiesis using marker combinations previously reported to be useful in this distinction. Our goal was to determine whether the previously reported ability of FCM to discriminate benign from dysplastic myelopoiesis could be reproduced when full datasets are objectively compared by this method.

MATERIALS AND METHODS

Data Retrieval and Preparation

Seventeen marrow samples diagnostic for MDS were retrieved from the clinical FCM archive of the Medical College of Wisconsin over a 3-year period. Twenty-five samples of non-neoplastic bone marrow (drawn either for staging of nonmyeloid neoplasms or for abnormalities of peripheral blood cell counts) from a comparable time interval were used as a control group. None of the cases in the control group showed evidence of lymphoma, dysplasia, or myeloid neoplasm. MDS cases included five refractory anemia with excess blasts (RAEB), five refractory cytopenia with multilineage dysplasia (RCMD), two refractory anemia with ring sideroblasts (RARS), two therapy related MDS (t-MDS), and three MDS, unclassified (MDS-U), based on 2008 World Health Organization diagnostic criteria (1). Data from cytogenetic analysis were available in 13 of the 17 MDS cases and 19 of the 25 non-neoplastic cases. For the specific comparison of cases with and without morphologic granulocyte dysplasia, 14 of the 17 MDS cases were available for morphologic reassessment and were graded for the presence or absence of morphologic dysgranulopoiesis in at least 10% of granulocyte precursors based on WHO 2008 criteria (1).

Because the FINE algorithm uses single high-dimensional datasets, we selected cases for analysis based on availability of data from a single tube that contained markers previously reported as useful in the analysis of dysgranulopoiesis in MDS (3,4,10–12): CD11b, CD16, CD45, and CD56. When available for a given case, we also separately analyzed a single tube that contained CD13, CD16, HLA-DR, and CD45, a combination also previously reported as useful in the analysis of the myeloid compartment in MDS (2–4,12). Each tube included analysis of forward and side angle light scatter for a total of six measurements per cell per assay and the creation of a six-dimensional probability density function for analysis in each case. All antibody reagents were obtained from Becton-Dickinson (Franklin Lakes, NJ), with fluorescent conjugates as follows: for the first antibody combination, CD11b-allophycocyanin (APC), CD16-fluorescein isothiocyanate (FITC), CD45-peridinin chlorophyll protein (PerCP), and CD56-phycoerythrin (PE); and for the second antibody combination, CD16-FITC, CD13-PE, HLA-DR-PerCP, and CD45-APC. Heparin or EDTA-anticoagulated bone marrow samples were prepared via ammonium chloride lysis, and after routine staining and fixation, each sample was analyzed using a Becton-Dickinson FACSCalibur flow cytometer, and 30,000 cellular events were captured for each assay. Data were collected using Cell Quest acquisition software (Becton-Dickinson) and saved in standard fcs format.

For each analysis, CD45-negative events (nucleated red blood cells, debris, etc) were excluded. Then, analyses were performed on (1) datasets that included all nonlymphoid CD45-positive cellular events (with separate analyses restricted to cases with available cytogenetics data and cases reviewable for morphologic dysgranulopoiesis) and (2) datasets restricted to cellular events resembling blasts based on CD45 and side angle light scatter characteristics. The gated list mode data were then converted to tab-delimited text using WinMDI software (Scripps Research Institute, La Jolla, CA) or Winlist 6.0 (Verity Software House, Topsham, ME).

Fisher Information Nonparametric Embedding Analysis

Fisher information nonparametric embedding (FINE) is a method by which the information contained within FCM datasets can be objectively compared by treating each dataset as a high-dimensional probability density function (or "shape") rather than as a series of two-dimensional histograms. The differences in the nature of the high-dimensional distributions created by each *n*-parameter FCM dataset are measured using computational estimates of the Fisher information distance (a measure of the differences among probability density functions embedded on statistical manifolds). The underlying principles of FINE and information geometry are described

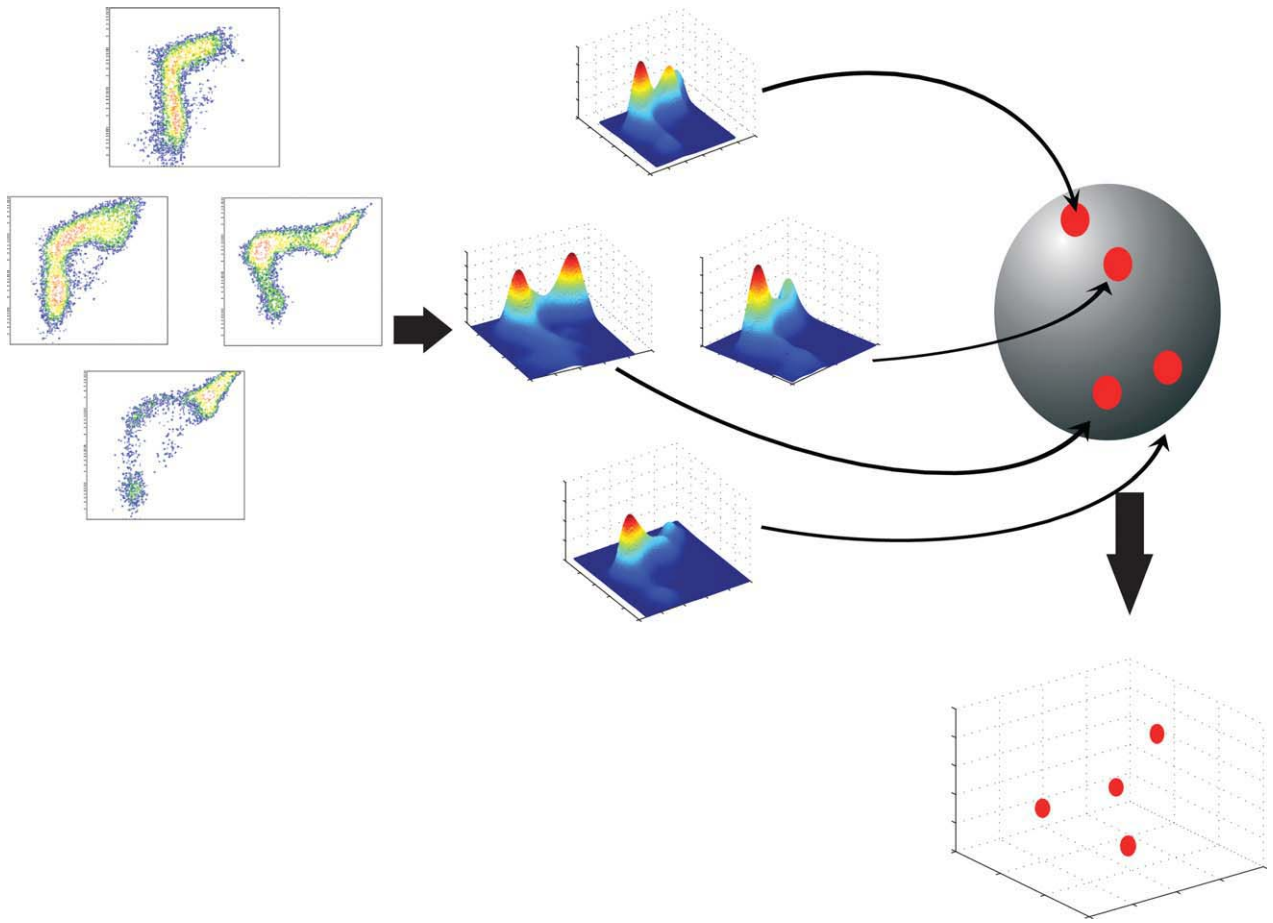

FIG. 1. Schematic overview of Fisher information nonparametric embedding (FINE). N-parameter event data from flow cytometry list mode files are converted into probability density functions (PDFs); the PDFs are embedded as points on a high-dimensional virtual construct known as a statistical manifold, and the differences in information contained within the PDFs are represented as distances along the statistical manifold (Fisher information distance) using a computational estimate; finally, the high-dimensional neighborhood map on the statistical manifold is reduced to a lower dimensional plot (two or three dimensions) for visualization. Details of the FINE method are provided in references (7,8).

in detail in earlier studies (7,8), and a schematic representation of the method is provided in Figure 1.

FINE analysis and IPCA analysis (see below) were performed in the MATLAB 2010b computing environment (MathWorks, Natick, MA) using the tab-delimited text data files as noted earlier. The MatLab FINE and IPCA codes are available at the University of Michigan Information Geometric Dimensionality Reduction (IGDR) Toolbox web page, https://tbayes.eecs.umich.edu/kmcarter/igdr/index.html.

Information Preserving Component Analysis

To illustrate and contrast the distinguishing characteristics of the six-parameter FCM datasets for a given case within the FINE plot of information distances, IPCA was performed to obtain information-optimized two-dimensional FCM histogram projections for each case. By necessity, FCM data must be projected via histograms of limited dimensionality (typically, two-dimensional projections) even though datasets are of higher dimensionality. Generally, these projections plot one measured parameter (either fluorescence intensity of a surface marker or a light scatter characteristic) against another measured parameter—in essence a "head-on" projection of the six-dimensional "object" in the case of six-parameter FCM. IPCA derives an information-optimized (in essence rotated) 2D projection of the six-dimensional dataset. The optimization is based upon determining the 2D projection that will best preserve the information distances between each case, as determined by the FINE algorithm. IPCA is explained in more detail in a previous publication (9) and involves minimization of the loss of information in the conversion from six-dimensional to two-dimensional distributions, based on the mathematical principle of gradient descent.

FINE and IPCA are both data visualization methods that are based on preserving information in the full high-dimensional FCM datasets. The fundamental difference between them is that FINE is used for visualizing a population of subjects, representing each subject's dataset as a point in two or three dimensional space, while IPCA is used for visualizing the optimal shape of each subject's dataset.

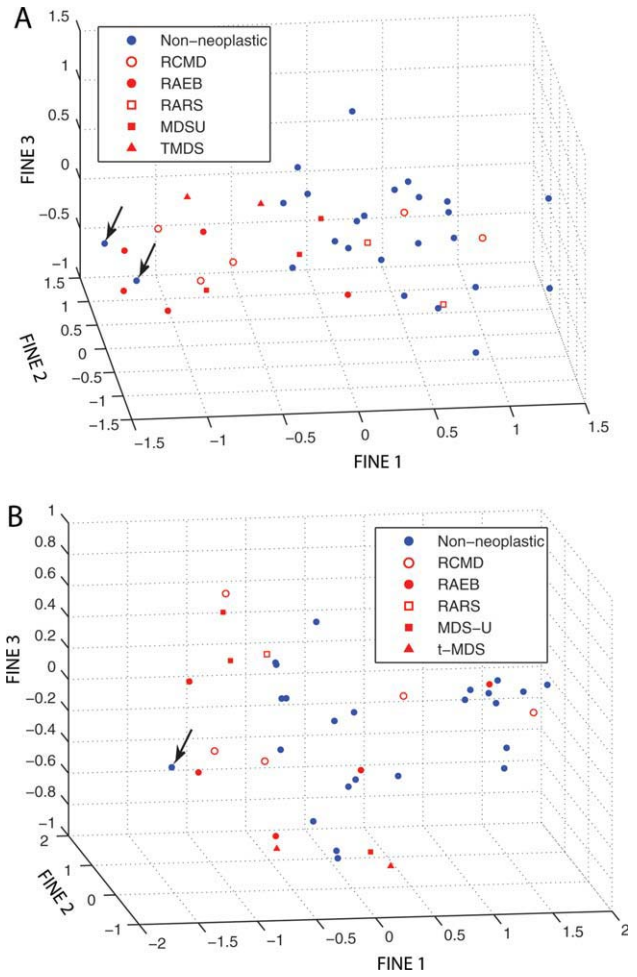

FIG. 2. Three-dimensional Fisher information nonparametric embedding (FINE) plots illustrating information distances between cases by analysis of flow cytometry datasets as single high-dimensional distributions. Panel A is based on a four-color (six-parameter) combination of forward scatter (FS)/side scatter (SS)/CD16/CD56/CD45/CD11b. Panel B is based on a four-color (six-parameter) combination of FS/SS/CD16/CD13/HLA-DR/CD45. Arrows indicate benign bone marrow samples located in a part of the FINE plot dominated by high-grade myelodysplastic syndromes. One such case was not available for the marker combination illustrated in panel B. The axes (FINE 1, FINE 2, and FINE 3) are derived components that represent computational estimates of relative Fisher information distance between each analyzed case. See text for additional comments. RCMD, refractory cytopenia with multilineage dysplasia; RAEB, refractory anemia with excess blasts; RARS, refractory anemia with ring sideroblasts; MDSU, myelodysplastic syndrome, unclassified; TMDS, therapy-related myelodysplastic syndrome.

RESULTS

For the analysis of nonlymphoid CD45 positive cells, three-dimensional projections of information distance maps between individual cases for each of the two four-color (six-parameter) marker combinations are shown in Figure 2. Exact distances varied somewhat depending upon the mathematical information distance measurement used [the data shown use the Hellinger distance (7)], but the trends were consistent and are accurately depicted in Figure 2. Although there was some overlap between MDS and non-neoplastic cases, both marker combinations revealed a region of FINE space dominated by MDS cases. Cases of "high grade" MDS (RAEB and RCMD) tended to occupy the areas of greatest information distance from non-neoplastic cases. Cases of RARS and subsets of each of the other MDS categories were essentially indistinguishable from non-neoplastic cases. This general pattern was reproduced through various iterations of the FINE algorithm using different information distance measurements and different nearest neighbor parameters for constructing the 2D and 3D FINE plots (data not shown).

The optimized two-dimensional projection matrices, derived via IPCA, are shown for each of the two marker combinations in Tables 1 and 2, respectively. Figure 3 also illustrates the patterns of CD11b versus CD16 histograms (projections), and optimized 2D projections derived via IPCA, for representative subsets of cases relative to their positions on the 3D FINE plots, in order to qualitatively illustrate the differences among cases that are represented by the FINE information distances. The axes on the IPCA plots are components derived from weighted contributions of each of the measured markers for each assay. For the CD16/CD56/CD45/CD11b combination, the *X* axis roughly corresponds to the difference between forward angle light scatter and the combined contribution of CD11b and CD16 expression, while the *Y* axis is more heavily weighted to forward and side angle light scatter characteristics. For the CD16/CD13/HLADR/CD45 combination, the *X* axis is heavily weighted toward the intensity of CD45 expression, while the *Y* axis is weighted toward a roughly equal contribution of CD16 intensity and side angle light scatter. The IPCA algorithm determined that these were the weightings that would best convey the differences among cases when the six markers in each case were projected into two dimensions.

Table 1
Projection Matrix for the Optimal Two-Dimensional Projection of Six-Parameter Data in Each Individual Case from a Single Tube Including Forward Angle Light Scatter, Side Angle Light Scatter, CD16, CD56, CD45, and CD11b

	Forward scatter	Side scatter	CD16	CD56	CD45	CD11b
X axis	0.6417	0.5812	0.3001	0.1457	0.2169	0.3034
Y axis	0.4468	0.0568	−0.6585	0.2067	0.0821	−0.5604

The numbers in each cell of the table indicate the relative weighting of the individual parameter in generating the two-dimensional projection for a given case that is most likely to preserve the information distances observed on FINE analysis. The information in this matrix is represented graphically in Figure 3.

Table 2
Projection Matrix for the Optimal Two-Dimensional Projection of Six-Parameter Data in Each Individual Case from a Single Tube Including Forward Angle Light Scatter, Side Angle Light Scatter, CD16, CD13, HLA-DR, and CD45

	Forward scatter	Side scatter	CD16	CD13	HLA-DR	CD45
X axis	0.0760	−0.0961	−0.1309	−0.0067	0.0922	0.9794
Y axis	0.0510	0.6435	0.6761	0.1530	−0.2694	−0.1739

See legends to Table 1 and Figure 3 for additional explanation.

Of note are two outliers (highlighted by arrows in Fig. 2). Two non-neoplastic marrow samples were located a considerable distance from other non-neoplastic cases on the FINE map using the CD16/CD45/CD56/CD11b combination. Only one of these two cases was present in the CD45/CD13/HLA-DR/CD16 combination, but again, it was located in the MDS region. One of these two cases was a marrow sample taken from a patient in early recovery from transient agranulocytosis, thought to be medication-related. The marrow in this patient's case showed a marked shift to immaturity ("maturation arrest") with eventual recovery of peripheral counts. The other of these two cases was taken from a patient with a history of heart transplant and immunosuppression therapy, who presented with a leukemoid reaction in which his peripheral leukocyte count reached over 100×10^9/l. This patient's bone marrow sample showed a marked granulocytic predominance with shift to immaturity. This patient was ultimately diagnosed with an infectious cause for the leukocytosis, and the leukocyte count corrected to 5×10^9/l within days after the bone marrow biopsy.

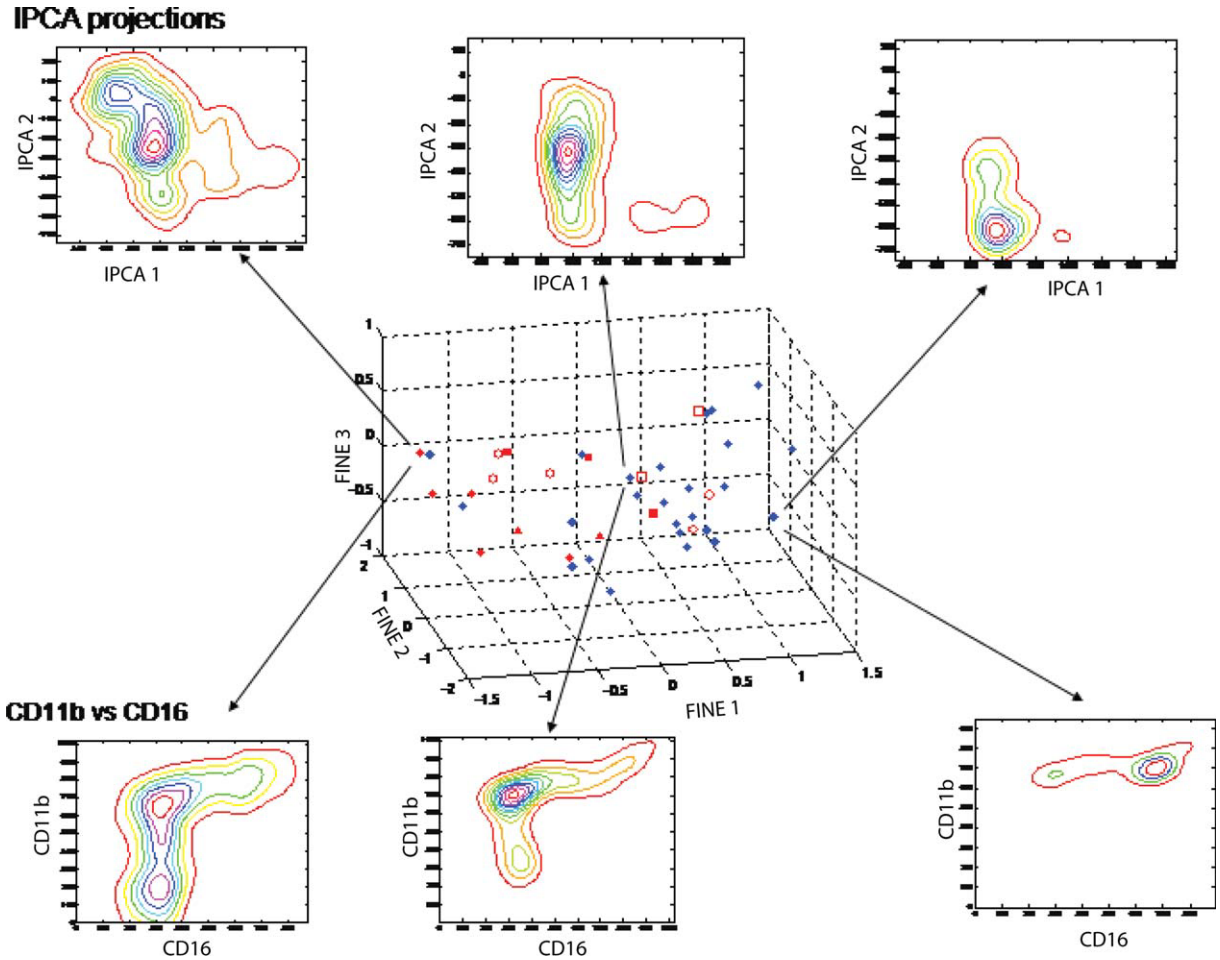

FIG. 3. FINE plot based on the FS/SS/CD16/CD56/CD45/CD11b combination, with overlay of two-dimensional flow cytometry contour plots for each case. The lower panel represents CD11b versus CD16. The upper panel represents an optimized rotated two-dimensional projection of the six-parameter (four-color) dataset, with optimization based on information-preserving component analysis (IPCA). See text for additional explanation, including the components of the IPCA axes. 2D plots are shown for a case located in the myelodysplastic region of the plot (left), the centroid region of the plot (center), and the nonneoplastic region of the plot (right). Note the difference in character of 2D projections between cases that are located in different regions of the FINE plot.

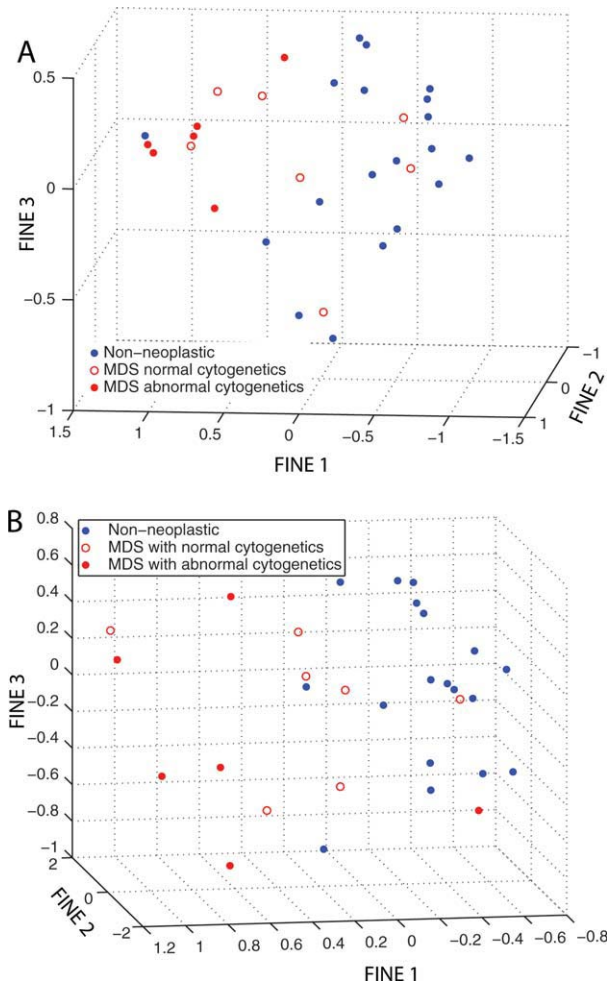

FIG. 4. Three-dimensional FINE plots limited to analysis of cases with available cytogenetic analysis data. Panel A shows the results of the FS/SS/CD16/CD56/CD45/CD11b combination, and panel B shows the results of the FS/SS/CD16/CD13/HLA-DR/CD45 combination. See text for more detail of cytogenetic results.

A separate analysis restricted to cases for which cytogenetic data were available is depicted in Figure 4. Seven of 13 MDS cases with available cytogenetics data showed clonal cytogenetic abnormalities, as follows: two cases of deletion 7q (one detected by fluorescence in situ hybridization only) and one case each of isochromosome X, deletion 20q, translocation t(3;5), trisomy 8, and a complex karyotype with numerous abnormalities. There was an area of FINE space concentrated toward MDS cases with clonal cytogenetic abnormalties, and this finding was apparent in datasets from each of the two marker combinations.

A separate analysis comparing cases with and without morphologic dysgranolopoiesis is depicted in Figure 5. As noted in the Methods section, 14 total MDS cases were available for this analysis (13 cases for the CD16/CD13/HLA-DR/CD45 marker combination). Each of these analyses revealed that MDS cases with morphologic dysplasia showed greater separation from benign cases than did MDS cases without morphologic dysgranulopoiesis.

When the analysis was restricted only to blast populations in all cases, there was no clear separation of MDS from non-neoplastic cases (data not shown). However, this analysis was hindered by low event numbers in many samples that may have influenced the quality of the probability density functions created from the limited datasets. Analysis of myeloblasts for cases that included a minimum of 1,000 captured events in that region revealed apparent separation of the three remaining nonneoplastic cases from all but one of the remaining MDS cases in the CD16/CD11b/CD56/CD45 analysis; but, again, this analysis was limited by the low number of cases meeting this threshold for analysis (Fig. 6). The CD16/CD13/HLA-DR/CD45 analysis using blasts only (for cases with a minimum of 1,000 events in the blast region) showed no clear pattern, but only included nine cases total including two non-neoplastic cases.

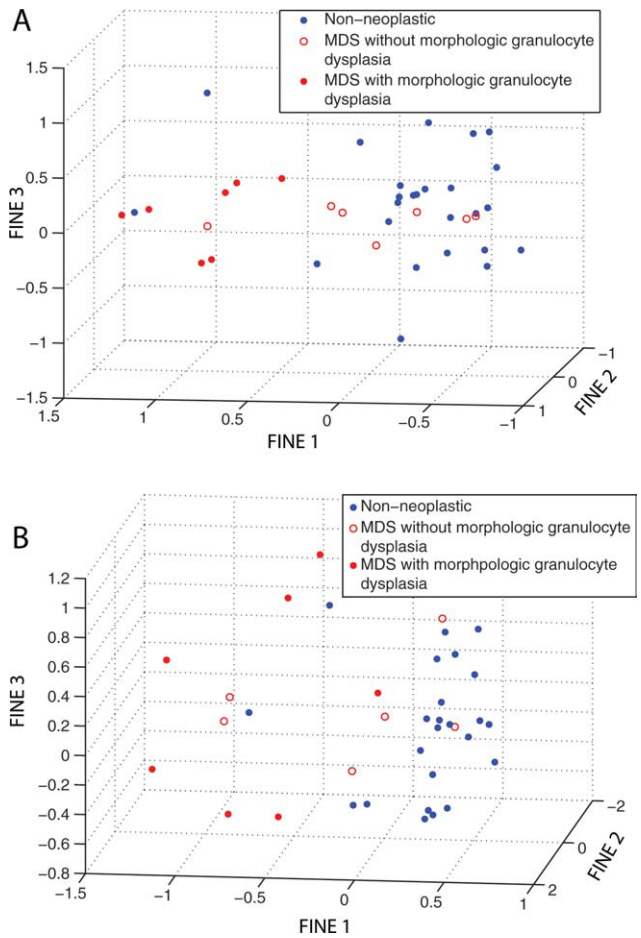

FIG. 5. FINE plot comparing MDS cases with and without morphologic dysgranulopoiesis in comparison with benign cases. MDS cases with morphologic dysgranulopoiesis show clearer separation from benign cases than do MDS cases without morphologic dysgranulopoiesis. Panel A represents the FS/SS/CD16/CD56/CD45/CD11b marker combination. Panel B represents the FS/SS/CD16/CD13/HLA-DR/CD45 combination.

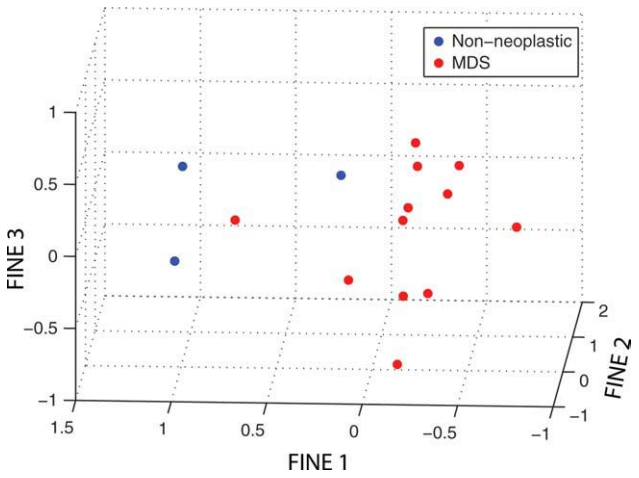

FIG. 6. FINE plot based on the FS/SS/CD16/CD56/CD45/CD11b combination for analysis limited to events resembling myeloblasts by CD45 and side angle light scatter characteristics, for which at least 1,000 event were captured in that region. Two of three nonneoplastic cases segregate from MDS cases; however, the analysis is hindered by the low number of cases and by the fact that the marker combinations were not optimized for blast immunophenotyping.

Myeloblasts as a percentage of CD45 positive events ranged in the MDS group from 0.3% to 9.2%. The MDS cases in the outermost portion of the FINE plot had between 5 and 9% myeloblasts; however, other cases in the MDS cluster had as few as 0.3% blasts.

DISCUSSION

Documentation of abnormal, imbalanced, or dyssynchronous myeloid maturation patterns has been a key focus of most studies on the use of flow cytometric immunophenotyping in the diagnosis of MDS (2–5,13–15). This type of analysis generally has a substantial subjective component, requiring an individual user to weigh qualitatively the expression patterns among several markers, to weigh the importance of maturational shifts or myeloblast increases, and to render conclusions regarding the sometimes subtle deviations and between observed and expected patterns yielded by these marker combinations on examination of serial 2D histograms.

Previous studies have demonstrated characteristic patterns of marker acquisition and expression in normal granulocyte maturation. For example, a 2D histogram of CD11b versus CD16 shows a right-angle pattern of gradual acquisition of CD11b, followed by gradual acquisition of CD16. CD13 versus CD16 typically shows a characteristic "sickle" pattern (2–5,16,17). CD45 versus side angle light scatter shows a characteristic pattern of progressively increased CD45 intensity and diminishing side scatter between promyelocyte and neutrophil stages and so on. Deviations from these patterns (due to asynchronous marker acquisition or abnormal granulation patterns) can be demonstrated in cases of abnormal myelopoiesis. However, it is difficult to quantify the extent to which granulocyte maturation patterns diverge in benign versus dysplastic granulopoiesis or to set consistent thresholds for interpreting this divergence. As a result, previous studies vary in the level of detail or precision of subjective evaluation criteria.

In a typical FCM dataset, thousands of cells are analyzed, each for a set of several specific characteristics-2 light scatter characteristics and, in the case of 4-color FCM, the expression characteristics of four specific antigens. The magnitude of each measured characteristic can be thought of as a spatial coordinate and each analyzed cell can be thought of as occupying a single point in high dimensional space (six-dimensional space in the case of four-color FCM) (18–21). The aggregate of these points then forms in essence a six-dimensional object. A typical FCM interpretation involves examining serial two-dimensional projections (histograms) of each of these six-dimensional objects. In contrast to this type of hierarchical two-dimensional analysis, information geometry—a field of information theory based largely on the work of Amari (22)—allows for the objective comparison of differences in information contained within high-dimensional probability density functions, of the type formed by routine clinical FCM data. This approach quantifies the differences in the shape and density of the high-dimensional "objects" formed by FCM datasets (7).

By applying the principles of information geometry in the current study, we corroborated observations that have been made in previous studies involving subjective interpretation of myeloid immunophenotypes, including substantial overlap in the immunophenotypic signatures between benign granulopoiesis and the granulopoiesis of low grade MDS and the tendency of higher grade MDS (RAEB and RCMD) to separate immunophenotypically from benign cases (5,6,15). This finding was further validated by the observation that MDS cases with morphologic dysgranulopoiesis showed greater separation from benign cases than did MDS cases without morphologic dysgranulopoiesis. We also noted a greater separation from normal in cases of MDS that were documented to harbor clonal cytogenetic abnormalities.

There are two important caveats to these observations. First, a small subset of the high-grade MDS cases embedded indistinguishably within the population of benign cases. Second, two cases of left-shifted granulopoiesis embedded with the high-grade MDS cases. The latter observation highlights a pitfall observed in previous studies, in which "dysplastic" maturation patterns (particularly abnormalities in the acquisition of CD11b, CD13, and CD16) were observed in regenerating or reactive bone marrow samples (13,15). The relative heterogeneity of the non-neoplastic cluster of samples in this study is likely due to the expected variability in normal granulopoiesis among samples. Although the sequence of marker acquisition in physiologic granulopoiesis is stable across different individuals, shifts in maturation along these sequences are common (17). We deliberately avoided excluding potentially "reactive" marrow samples in the non-neoplastic group in this study for two reasons. First, given the marked variability in

environmental stimuli affecting granulopoiesis, it would be difficult to confirm that a given sample was truly normal or that normal cases would not include some degree of reactive change in granulocyte maturation. Second, the day-to-day clinical exercise of elucidating dysplastic myelopoiesis by FCM includes the broad array of non-neoplastic changes that can affect myeloid precursors. Differences in immunophenotype detected by our analysis were likely influenced by blast percentage, with the MDS cases in the outermost regions of the information distance plot showing blast percentages between 5 and 9%. However, the position on the information distance plots was not entirely linked to blasts, because cases with fewer than 1% blasts were also located in regions near cases with increased blasts. Data restricted only to blast populations showed a trend toward separating MDS from non-neoplastic cases, but this analysis was hindered by the low number of cases with sufficient blasts for meaningful analysis.

Part of the current study involved using the information distances between cases to determine optimized 2D projections for the display of flow cytometric findings for each individual case via IPCA. It is interesting that, when analyzing the CD16/CD56/CD45/CD11b combination, the derived components that constitute the axes of each IPCA plot were weighted toward the combined contribution of CD11b and CD16 on one axis and to light scatter characteristics on the other axis. CD11b and CD16 have been used together in several studies to characterize the spectrum of myeloid maturation in blood and bone marrow, and abnormalities in light scatter characteristics have also been used to detect abnormal patterns of cytoplasmic granulation in myelodysplastic bone marrow (3,4,16). IPCA projections of the CD16/CD13/HLA-DR/CD45 combination were heavily weighted toward CD45 on one axis and to the combined contribution of CD16 and side angle light scatter on the other axis.

The observations in the current study are limited by the fact that previous studies have evaluated not just granulopoiesis, but also abnormalities in erythropoiesis and megakaryopoiesis in the flow cytometric analysis of MDS (4,23). Additional studies applying information geometry to marker combinations designed to evaluate maturation patterns in these other cell lineages would be useful to further test the conclusions of the current study. The current study is also somewhat limited by the use of only four colors (six parameters) in each analysis, reflective of the constraints of archived clinical FCM data given the relatively recent expansion of polychromatic FCM in routine clinical use. An optimal approach may have included several additional markers relevant to stages of myeloid maturation, including CD34, CD117, and others, but, by design, this study focused on markers expressed to varying degrees throughout myeloid maturation that are commonly used in myelodysplasia assessment. As clinical-grade instruments are expanding to 8- and 10-color analysis routinely, future studies should also focus on the prospective application of information geometry-based cytomic profiling to higher dimensional FCM datasets to further clarify the observations noted in this retrospective study (24). In addition, our analyses are subject to potential variability in staining or acquisition between analyses in the day-to-day practice of clinical cytometry; however, this is a limitation that also applies to a substantial majority of published studies evaluating flow cytometric immunophenotyping in clinical diagnostics.

This study illustrates the application of a novel cytomic profiling tool to the assessment of granulocyte maturation patterns by FCM. When FCM datasets are objectively compared as single high-dimensional distributions, there are differences discernable between benign granulopoiesis and RAEB/RCMD. These differences are most evident in cases with overt morphologic dysgranulopoiesis or cases that harbor clonal cytogenetic abnormalities. However, some types of MDS overlap considerably with benign cases with respect to the immunophenotype of granulopoietic maturation. Furthermore, there is overlap between left shifted granulopoiesis and even high grade MDS, highlighting a potential pitfall noted in previous studies (3,15).

LITERATURE CITED

1. Brunning RD, Orazi A, Germing U, Le Beau MM, Porwit A, Baumann I, Vardiman JW, Hellstrom-Lindberg E. Myelodysplastic syndromes/neoplasms, overview. In: Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW, editors. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Lyon: International Agency for Research on Cancer; 2008. pp 87–107.
2. Kussick SJ, Fromm JR, Rossini A, Li Y, Chang A, Norwood TH, Wood BL. Four-color flow cytometry shows strong concordance with bone marrow morphology and cytogenetics in the evaluation for myelodysplasia. Am J Clin Pathol 2005;124:170–181.
3. Loken MR, van de Loosdrecht A, Ogata K, Orfao A, Wells DA. Flow cytometry in myelodysplastic syndromes: Report from a working conference. Leuk Res 2008;32:5–17.
4. Stetler-Stevenson M, Arthur DC, Jabbour N, Xie XY, Molldrem J, Barrett AJ, Venzon D, Rick ME. Diagnostic utility of flow cytometric immunophenotyping in myelodysplastic syndrome. Blood 2001;98:979–987.
5. Wells DA, Benesch M, Loken MR, Vallejo C, Myerson D, Leisenring WM, Deeg HJ. Myeloid and monocytic dyspoiesis as determined by flow cytometric scoring in myelodysplastic syndrome correlates with the IPSS and with outcome after hematopoietic stem cell transplantation. Blood 2003;102:394–403.
6. Wells DA, Ogata K. On flow cytometry in myelodysplastic syndromes, with caveats. Leuk Res 2008;32:209–210.
7. Carter KM, Raich R, Finn WG, Hero AO III. FINE: Fisher information nonparametric embedding. IEEE Trans Pattern Anal Mach Intell 2009;31:2093–2098.
8. Finn WG, Carter KM, Raich R, Stoolman LM, Hero AO. Analysis of clinical flow cytometric immunophenotyping data by clustering on statistical manifolds: Treating flow cytometry data as high-dimensional objects. Cytometry B Clin Cytom B 2009;76B:1–7.
9. Carter KM, Raich R, Finn WG, Hero AO. Information preserving component analysis: Data projections for flow cytometry analysis. IEEE J Select Topics Signal Process 2009;3:148–158.
10. Miller DT, Stelzer GT. Contributions of flow cytometry to the analysis of the myelodysplastic syndrome. Clin Lab Med 2001;21:811–828.
11. Truong F, Smith BR, Stachurski D, Cerny J, Medeiros LJ, Woda BA, Wang SA. The utility of flow cytometric immunophenotyping in cytopenic patients with a non-diagnostic bone marrow: A prospective study. Leuk Res 2009;33:1039–1046.
12. Bowen KL, Davis BH. Abnormal patterns of expression of CD16(FcR-III) and CD11b(CRIII) antigens by developing neutrophils in the bone marrow of patients with myelodysplastic syndrome. Lab Hematol 1997;3:292–298.

13. Loken MR, Wells DA. The role of flow cytometry in myelodysplastic syndromes. J Natl Compr Canc Netw 2008;6:935–941.
14. Maynadie M, Picard F, Husson B, Chatelain B, Cornet Y, Le Roux G, Campos L, Dromelet A, Lepelley P, Jouault H, et al. Immunophenotypic clustering of myelodysplastic syndromes. Blood 2002;100: 2349–2356.
15. Stachurski D, Smith BR, Pozdnyakova O, Andersen M, Xiao Z, Raza A, Woda BA, Wang SA. Flow cytometric analysis of myelomonocytic cells by a pattern recognition approach is sensitive and specific in diagnosing myelodysplastic syndrome and related marrow diseases: Emphasis on a global evaluation and recognition of diagnostic pitfalls. Leuk Res 2008;32:215–224.
16. Fujimoto H, Sakata T, Hamaguchi Y, Shiga S, Tohyama K, Ichiyama S, Wang FS, Houwen B. Flow cytometric method for enumeration and classification of reactive immature granulocyte populations. Cytometry 2000;42:371–378.
17. van Lochem EG, van der Velden VH, Wind HK, te Marvelde JG, Westerdaal NA, van Dongen JJ Immunophenotypic differentiation patterns of normal hematopoiesis in human bone marrow: Reference patterns for age-related changes and disease-induced shifts. Cytometry B Clin Cytom B 2004;60B:1–13.
18. Roederer M, Hardy RR. Frequency difference gating: A multivariate method for identifying subsets that differ between samples. Cytometry 2001;45:56–64.
19. Roederer M, Moore W, Treister A, Hardy RR, Herzenberg LA. Probability binning comparison: A metric for quantitating multivariate distribution differences. Cytometry 2001;45:47–55.
20. Zamir E, Geiger B, Cohen N, Kam Z, Katz BZ. Resolving and classifying haematopoietic bone-marrow cell populations by multi-dimensional analysis of flow-cytometry data. Br J Haematol 2005;129: 420–431.
21. Zeng QT, Pratt JP, Pak J, Ravnic D, Huss H, Mentzer SJ. Feature-guided clustering of multi-dimensional flow cytometry datasets. J Biomed Inform 2007;40:325–331.
22. Amari S, Nagaoka H. Differential-Geometrical Methods in Statistics. Berlin: Springer-Verlage; 1990.
23. Della Porta MG, Malcovati L, Invernizzi R, Travaglino E, Pascutto C, Maffioli M, Galli A, Boggi S, Pietra D, Vanelli L, et al. Flow cytometry evaluation of erythroid dysplasia in patients with myelodysplastic syndrome. Leukemia 2006;20:549–555.
24. Wood B. 9-color and 10-color flow cytometry in the clinical laboratory. Arch Pathol Lab Med 2006;130:680–690.