Program Number: 230

To understand and detect abnormal states in any complex system requires detailed knowledge of the accompanying normal states. This knowledge of normal versus abnormal patterns is of particular importance in the discovery of abnormal cells in tissue specimens by the technology of cytometry. Typically, cytometry experts in the general area of hematopathology examine a set of bivariate dot plots derived from a specific reagent panel and look for the presence of cell populations that have unusual characteristics. This investigation is often guided by ancillary information about the patient's clinical history and the results from other tests. These experts know the normal expression patterns for each pair of markers and can often find small malignant populations at a sensitivity a few tenths of a percent.

Since the health care industry is being pressured to cut costs without loss of quality, automating expensive medical tests is likely to play an increasingly important role in the future. This study attempts to answer the question of whether a fully autonomous computational system can mathematically model the complex normal development of B-cells in bone marrow. Since B-cell malignancies are the most common form of leukemia and lymphoma, these results may have general applicability to all clinical cytometry laboratories.

This study will use Probability State Modeling (PSM) as the central computational engine because of its ability to scale well with number of measurements and its accuracy in accounting for population overlap due to inevitable errors in the measurement process. Thirteen different listmode files from "uninvolved" bone marrow specimens will be subjected to this automatic analysis and the results will be compared with expert analyses of the same data. After normal B-cells are modeled, the system detects events that for some reason are at the edges of the normal B-cell model's probability distribution and characterizes them.

Automated Modeling of Human B-Cell Progressions In Bone Marrow and Detection of Minimum Residual Disease

C. B. Bagwell[1], C. C. Stewart[2], B. Wood[3], and F. Preffer[4]

[1]Verity Software House, Topsham, Maine 04086 USA
[2]Emeritus Director, Laboratory of Flow Cytometry, Roswell Park Cancer Institute, Buffalo, NY 14063 USA
[3]University of Washington, Dept. of Laboratory Medicine, Seattle WA 98195 USA
[4]Harvard Medical School, Massachusetts General Hospital, Dept. of Pathology, Boston 02114 USA

1 Poster Narrative

In this study thirteen high-dimensional listmode files were used to test automatic analyses of normal human B-cell lineages in bone marrow. All thirteen of these files were reported out as being "uninvolved" for any malignancy involving B-cells. Table 1 (below) in the Panel 2 describes the markers, fluorochromes, and total number of events in each file. All the files were acquired on a BD LSRII using BD FACSDiva Software Version 4.0 and compensated for signal crossover.

The automation analysis routines (GemStone Version 1.50, Verity Software House) were designed to perform parameter name matching, making the system capable of full automation with a single template document even though markers were not always in the same position in the file and not always reported by the same fluorochrome.

Once a B-cell template document was read into the system, all of the thirteen files were processed with an algorithm defined by a template model document. The general logic of this template model is summarized in Panel 3. In all thirteen cases, the normal B-cell lineage was successfully modeled within one to three minutes. Two of the generated overlay plots that summarize all the marker correlations in the data are shown in Panel 2, bottom. Although the percentages in each stage and the marker intensities varied greatly from patient to patient, normal B-cells always displayed the same coordination of changes in marker intensities. As an example, when CD34 down-regulated there was a slight increase in CD45, CD19, and CD38 while there was a slight decrease in intensity for CD10. Other transitions in the B-cell lineage have similar patterns of reproducible coordination in markers.

All thirteen of the files were independently analyzed by four operators in different laboratories. The comparisons between the manual gating and automatic modeling approaches for the staging of the normal B-cell progressions are shown in Table 2, Panel 4. The manual gating (see Mean and SD columns in Table 2) and the automatic modeling (see Estimate column in Table2) estimates were found to be reasonably close to each other.

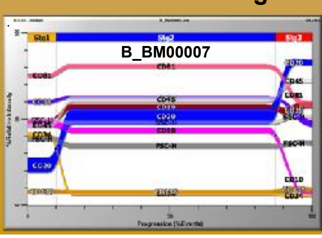
Once normal B-cells are modeled, it is possible to find small populations that for some reason are different than normal. Very small abnormal populations were introduced into data produced by the modeling system to simulate the presence of minimum residual disease (see Table 3, Panel 5 for phenotypes and percentages). The modeling Heat Map shows the presence of these small populations in each of the three stages of B-cell lineage (Panel 5). The normal B-cell lineage is shown below the Heat Map. The new TriCOM system, Panel 6, also shows these abnormal populations with a graphical depiction of their respective phenotypes. This file was sent to four operators in different laboratories to determine whether these three different abnormal populations could be detected with conventional gating methods. The results of their analyses are summarized in Table 4, Panel 5. In general, conventional gating methods often missed the presence of these abnormal populations. Also, the length of time necessary to analyze the data was an order of magnitude greater than the modeling system.

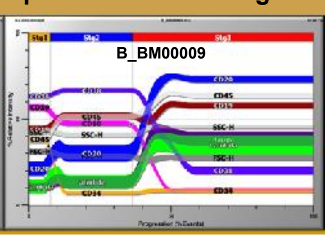
These results demonstrate that it will be feasible to automatically analyze diverse bone marrow specimens for the presence of very small aberrant populations that may indicate resistance to current therapeutic modalities.

2 Table 1: B-Cell Panels

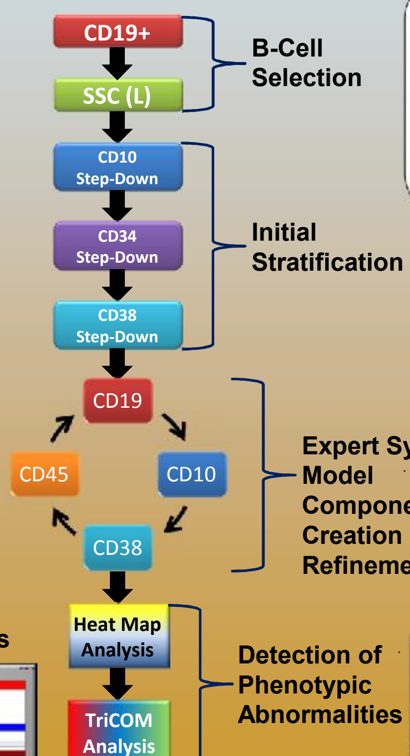
Files	Markers/Fluors									Total Events
B_BM00006	Kappa (FITC)	Lambda (PE)	CD19 (PE-TR)	CD34 (PerCP-Cy55)	CD20 (PE-Cy7)	CD45 (PB)	CD38 (A594)	CD10 (APC)	CD5 (APC-Cy7)	269,218
B_BM00007	CD81 (FITC)	CD10 (PE)	CD34 (PerCP-Cy55)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD138 (APC)	CD45 (APC-Cy7)		337097
B_BM00008	Kappa (FITC)	Lambda (PE)	CD34 (PerCP-Cy55)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-Cy7)		417911
B_BM00009	Kappa (FITC)	Lambda (PE)	CD34 (PerCP-Cy55)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-Cy7)		422432
B_BM00010	Kappa (FITC)	Lambda (PE)	CD34 (PerCP-Cy55)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-Cy7)		632595
B_BM00011	Kappa (FITC)	Lambda (PE)	CD34 (PerCP-Cy55)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-Cy7)		742320
B_BM00012	Kappa (FITC)	Lambda (PE)	CD34 (PerCP-Cy55)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-Cy7)		522838
B_BM00013	CD9 (FITC)	CD34 (PE)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-H7)			171218
B_BM00014	CD9 (FITC)	CD34 (PE)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-H7)			359633
B_BM00015	CD9 (FITC)	CD34 (PE)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-H7)			321185
B_BM00016	CD9 (FITC)	CD34 (PE)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-H7)			290688
B_BM00017	CD9 (FITC)	CD34 (PE)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-H7)			371037
B_BM00018	CD9 (FITC)	CD34 (PE)	CD19 (PE-Cy7)	CD20 (V450)	CD38 (A594)	CD10 (APC)	CD45 (APC-H7)			415403

GemStone Modeling Examples of B-Cell Lineages





3 Automated B-Cell Analysis Algorithm



4 Table 2: Manual Gating Versus Automatic Modeling Comparisons

	%B Cells *				%Stg1 **				%Stg2 **				%Stg3 **				%PC *			
	Gating		Modeling	Diff	Gating		Modeling	Diff	Gating		Modeling	Diff	Gating		Modeling	Diff	Gating		Modeling	Diff
Files	Mean	SD	Estimate	Est-Mean	Mean	SD	Estimate	Est-Mean	Mean	SD	Estimate	Est-Mean	Mean	SD	Estimate	Est-Mean	Mean	SD	Estimate	Est-Mean
B_BM_00006	12.63	0.12	12.08	0.55	8.82	2.03	11.13	2.32	53.19	4.98	42.60	-10.58	36.68	3.01	46.27	9.58	0.51	0.17	0.50	-0.01
B_BM_00007	11.52	0.15	10.95	0.58	9.36	4.55	8.73	-0.63	58.42	38.39	77.34	18.92	29.06	38.16	13.93	-15.13	0.52	0.12	0.31	-0.21
B_BM_00008	2.99	0.16	2.65	0.33	55.52	0.42	51.79	-3.73	37.54	2.42	45.81	8.27	6.06	2.18	2.41	-3.65	3.40	1.64	1.40	-2.00
B_BM_00009	3.42	0.34	2.87	0.55	8.51	0.81	7.84	-0.67	30.80	2.59	29.35	-1.45	58.93	0.67	62.80	3.87	6.52	5.24	5.13	-1.40
B_BM_00010	5.20	0.23	5.10	0.10	14.15	3.48	10.84	-3.32	71.03	4.35	76.51	5.47	12.72	1.49	12.66	-0.06	0.57	0.39	0.18	-0.39
B_BM_00011	4.11	0.14	3.62	0.48	23.75	0.67	21.54	-2.21	35.82	0.94	34.98	-0.84	38.83	1.11	43.48	4.64	0.32	0.23	0.35	0.03
B_BM_00012	4.94	0.09	4.74	0.19	5.43	0.30	3.98	-1.44	21.20	0.45	19.57	-1.63	71.86	0.93	76.44	4.58	0.37	0.12	0.51	0.14
B_BM_00013	21.04	0.45	19.99	1.05	5.38	0.71	3.69	-1.69	90.53	1.58	94.34	3.81	3.45	0.74	1.97	-1.49	0.04	0.03	0.00	-0.04
B_BM_00014	9.37	0.11	8.60	0.77	8.86	0.61	7.61	-1.25	35.57	0.80	31.88	-3.69	54.88	0.75	60.51	5.63	0.58	0.24	0.08	-0.50
B_BM_00015	3.21	0.23	2.87	0.34	41.92	2.96	48.19	6.27	32.53	3.41	31.60	-0.93	16.28	2.45	20.21	3.93	1.99	0.43	1.04	-0.94
B_BM_00016	0.70	0.04	0.69	0.01	53.15	4.56	74.38	21.23	24.13	0.90	19.07	-5.06	7.07	2.39	6.55	-0.52	1.29	0.79	1.24	-0.04
B_BM_00017	1.81	0.11	1.61	0.21	78.86	2.77	75.51	-3.35	15.48	1.96	23.82	8.34	4.64	1.18	0.67	-3.97	1.94	0.60	1.34	-0.60
B_BM_00018	3.16	0.02	2.96	0.19	42.52	1.46	38.10	-4.42	49.68	1.35	52.29	2.60	6.23	0.27	9.61	3.39	0.17	0.28	0.01	-0.16
Averages	6.47	0.17	6.06	0.41	27.40	1.95	27.95	0.55	42.76	4.93	44.55	1.79	26.67	4.26	27.50	0.83	1.40	0.79	0.93	-0.47

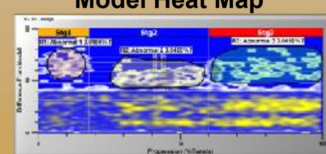
*: Percent of total events in file
**: Percent of classified B-cells

5 Table 3: Abnormal Populations

	Phenotypes								
Population	CD19	SSC	CD10	CD34	CD38	CD45	# Abnormalities	# Events	Percent
Abnormal 1	Normal	Normal	High	High	Low	Normal	3	57	0.03
Abnormal 2	Normal	Normal	Low	Normal	High	Normal	2	135	0.07
Abnormal 3	Normal	High	High	Normal	Normal	High	3	100	0.05

Total B-Cells
200300

Model Heat Map



Normal B-Cell Progression

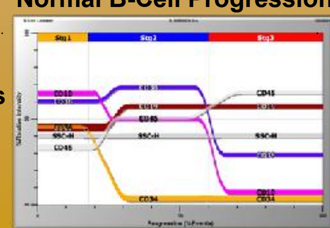
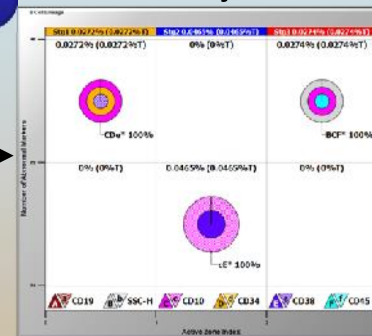


Table 4: Gating Results

Operators	# Correct	#Incorrect/Missing	Time (Hrs)
1	2	1	0.5
2	1	2	1
3	0	3	0.2

Four operators were given the above data to analyze for abnormal phenotypes using conventional listmode analysis. In general, most operators were not able to appreciate the small clusters of abnormal phenotypes that range in percentages from 0.02% to 0.05% (see Heat Map above and TriCOM display in Panel 6.

6 TriCOM System



How to read the above TriCOM graphic

The x-axis depicts the stages of the progression. In this B-cell example there are three stages. The y-axis quantifies the number of abnormalities in each phenotype (see Table 3). The key at the bottom shows how to interpret the phenotypes of the abnormal populations. For example, solid blue (E) represents higher than normal CD38; whereas, low density blue (e) represents lower than normal CD38. Reading from the outside ring to the center, the top-left pie chart shows that there is an abnormal population that is high for CD10 and CD34, but low for CD38. The percentages above the pie chart shows the percent of Stg1 events with three abnormalities. If there were more phenotypes discovered with three abnormalities, they would be shown as slices of the pie and their relative percentage shown next to each segment. This one graph allows the inspection of all abnormal phenotypes that may be present in a sample.

7 Discussion and Summary

The clinical utility of cytometry is likely to change from manual and subjective gating methods to automatic and objective methods such as shown in this study. The FDA is likely to require this kind of automation in routine clinical testing in the very near future.

In this study it was demonstrated that it was feasible to automatically model B-cell lineages defined by numerous cytometric measurements. It was also shown that once the normal components of bone marrow were mathematically described, it was possible to visualize very small aberrant populations by the GemStone Heat Map and TriCOM systems.

This study represents some early steps in leveraging modeling systems such as probability state modeling in the general area of minimum residual disease detection.

Thanks to the entire Verity team for making this study possible (Ben Hunsberger, Don Herbert, Mark Munson, and Chris Bray). Also, thanks to Margaret Inokuma at BD Biosciences for her help analyzing the files.