Laboratory Hematology 8:134-142

ISLH
Official Publication

Discrimination of Chronic Lymphocytic Leukemia of B-Cell Type by Computerized Analysis of 3-Color Flow Cytometric Immunophenotypes of Bone Marrow Aspirates and Peripheral Blood

G. VALET,[1] M. ARLAND,[2] A. FRANKE,[3] CH. KAHL,[4] H.-G. HÖFFKES[2]

[1]*Max-Planck-Institute of Biochemistry, Martinsried;* [2]*Klinikum Fulda, Medizinische Klinik III, Fulda;*
[3]*Division of Hematology/Oncology, Department of Medicine, Otto-von-Guericke-University of Magdeburg;*
[4]*Division of Hematology/Oncology, Department of Medicine, University of Rostock, Germany*

ABSTRACT

Restriction of κ- and λ-chain expression on CD19/CD5 double-positive bone marrow and peripheral blood lymphoid cells discriminates between malignant lymphoma and reactive lymphoid cells. Although this discrimination is unproblematic for high light-chain expression, difficulties arise for low-level expression or when a few malignant cells are dispersed in a majority of normal or reactive lymphocytes. This study investigated the suitability of the κ/CD19/5, λ/CD19/5, CD45/14/20, CD4/8/3 antibody panel for automated malignancy discrimination. Normal/κ B-cell chronic lymphocytic leukemia (B-CLL)/λ B-CLL bone marrow aspirate (BMA) samples were identified with predictive values of 100.0%/100.0%/96.4%. Normal peripheral blood leukocytes (PBL) can be used as a reference for κ/λ B-CLL BMA (100.0%/92.6%/92.9%) as a simplification. PBL samples were classified with slightly lower predictive values of 95.0%/96.9%/95.0%. The unknown embedded test set of clearly recognizable κ and λ B-CLL PBL and BMA samples was identified with predictive values of 100%/100% and 90.0%/100%. Manually unclassifiable BMA and PBL samples of 5 patients were classified by the CLASSIF1 algorithm (http://www.biochem.mpg.de/valet/classif1.html) in a logically consistent way, showing the usefulness of automated classification in otherwise difficult cases. In 2 patients, PBL samples were recognized as malignant, but no distinction between κ and λ B-CLL was possible. Data pattern analysis also detects differences between κ PBL and κ BMA cells in B-CLL patients with predictive values of 98.3%/100.0% and between λ PBL and λ BMA with values of 92.6%/96.4%. Furthermore, an automated BMA quality assessment can be performed with predictive values of 98.3%/100.0% for PBL/BMA discrimination. Together, these results show the interesting potential of medical bioinformatics data mining for the automated classification of complex clinical multiparameter flow cytometry data. *Lab Hematol.* 2002;8:134-142

KEY WORDS: Chronic lymphocytic leukemia · Immunophenotype · Flow cytometry · Cytomics · Medical bioinformatics · Data mining

Correspondence and reprint requests: H.-G. Höffkes, MD, Medizinische Klinik III, Klinikum Fulda, Pacelliallee 4, D-36043 Fulda, Germany (e-mail: hoeffkes.medIII@klinikum-fulda.de).

The data analysis for this study was done in accordance with the principles developed by the Data Analysis Task Force of the European Working Group for Clinical Cell Analysis (EU Concerted Action BMH4-CT97-2611[DG12-SSMI], http://www.ewgcca.org).

Received March 4, 2002; accepted April 9, 2002

INTRODUCTION

Flow cytometry (FCM) is a help in the distinction of malignant lymphoma cells from reactive lymphoid cells such

as in hyperplastic lymph nodes, although FCM-determined antigen expression does not directly correlate with classification by visual morphology [1,2]. Malignant lymphomas are diagnosed by FCM by light-chain restriction, that is, by an imbalance of cell surface κ- and λ-chains. Difficulties may arise in the case of infrequent lymphoma cells, eg, in an abundance of reactive cells. The detection of aberrant B-cell phenotypes by immunohistochemistry, as an alternative, is not reliable, and some antigens are difficult to detect by immunohistochemistry [3-5].

The development of self-learning multiparameter data classifiers substantially facilitates the classification of cytometric, clinical, and/or experimental multiparameter data in medicine. Large data sets from a variety of sources are rapidly processed, and laboratory and instrument-independent classification is possible [6-15]. This development is particularly useful for FCM, in which the simultaneous analysis of several cell markers generates substantial amounts of information.

In this study, the characteristics of lymphocytic, monocytic, and granulocytic or blastic cells in bone marrow and peripheral blood leukocytes (PBL) of patients with B-cell chronic lymphocytic leukemia (B-CLL) and age-matched healthy individuals were evaluated from immunophenotyping FCM list mode files. The numeric results were introduced into databases and classified by a nonparametric data pattern classifier [8,9]. Expert confirmed histopathological diagnosis of bone marrow trephines as "normal" or "B-CLL" served as "truth" for the self-learning process [16]. The resulting classifiers were tested for robustness against unknown patient samples deriving either from clearly normal and abnormal cases or from doubtful cases.

PATIENTS, MATERIALS, AND METHODS

Patients

Age-matched healthy volunteers and patients with B-CLL were included in this study. After informed consent had been given, heparinized bone marrow aspirates (BMA) were drawn from adult volunteers selected as bone marrow donors for a compatible or semicompatible allograft. The age distribution of these individuals ranged from 20 to 45 years, with an average of 29 years. None of the donors was taking medication except for birth control pills. Diagnostic specimens included fresh BMA and peripheral blood (PBL). Bone marrow trephines were infiltrated to variant degrees by lymphoma cells as confirmed by pathological and cytological examinations as well as by flow cytometric immunophenotyping. Samples were divided into 2 analysis groups: (1) the learning set contained BMA (22 healthy volunteers and 53 patients) as well as peripheral blood (58 healthy volunteers and 72 patients), and (2) the unknown test sets of (*a*) clear cases (including 8 normal samples, 9 κ and 3 λ B-CLL BMA, and 5 κ and 5 λ B-CLL PBL) and (*b*) unclear cases (4 BMA and 5 PBL samples from a total of 5 patients).

Histopathology

Fresh BMA were fixed in 5% formalin. Sections were stained with hematoxylin and eosin, periodic acid Schiff with and without prior diastase digestion, and reticulin. Selected samples were stained for nonspecific esterase and acid phosphatase in frozen sections. Each biopsy specimen was examined by 2 independent experienced hematopathologists. The histological subclassification was done according to the World Health Organization classification [16].

Immunophenotyping

Leukocyte concentrations in peripheral blood and bone marrow samples were adjusted with phosphate-buffered saline (PBS) to 5×10^6 cells/mL. Monoclonal antibodies (10-20 μL/100 μL diluted blood) were added in pretitered concentrations followed by vortexing and a 30-minute incubation at 0°C with vortexing every 10 minutes. A quantity of 2 mL Ortho-Lyse (Ortho, Heidelberg, Germany) was added to the samples with immediate vortexing followed by a 10-minute incubation at 4°C for erythrocyte lysis. The samples were then washed twice with 4 ml PBS by centrifugation for 5 minutes at 400*g*. The supernatant was discarded and the sediment resuspended in 2 mL PBS and 0.4 mL PBS/0.5% bovine serum albumin/0.1% Na azide and kept in the dark on ice until the flow cytometric measurement.

Antibody Panel

The following agents were used: fluorescein isothiocyanate (FITC)–coupled monoclonal CD45 (T29/33, immunoglobulin G1 [IgG1] isotype; DAKO Diagnostika, Hamburg, Germany), κ (rabbit antihuman F(ab′)2, polyclonal; DAKO), λ (rabbit antihuman F(ab′)2, polyclonal; DAKO), and CD4 (SK3, IgG1 isotype, Becton Dickinson, Heidelberg, Germany); R-phycoerythrin (PE)–coupled CD8 (SK1, IgG1 isotype, Becton Dickinson), CD14 (TUEK 4, IgG1 isotype; DAKO), and CD19 (4G7, IgG1 isotype; Becton Dickinson); peridinin chlorophyll a protein (PerCP) CD3 (SK7, IgG1 isotype, Becton Dickinson) and CD20 (L27, IgG1 isotype; Becton Dickinson); and phycoerythrin/cyanine 5 (PE-Cy5)–coupled CD5 (5D7, IgG1 isotype; Caltag Laboratories, San Francisco, CA). The combinations CD45/14/20, CD4/8/3, κ/CD19/5, and λ/CD19/5 were analyzed by multiparametric FCM for all samples drawn.

Flow Cytometry

Cell samples were processed within 2 hours after specimen collection. Analysis was performed on a FACScan (Becton Dickinson) analytical flow cytometer. List mode data were acquired by Lysis II software (Becton Dickinson). The cellular forward scatter (FSC) and sideward scatter (SSC) signals as well as the fluorescence of cell membrane–bound FITC-, PE-, PerCP-, or PE-Cy5–labeled antibody were determined following illumination of the cells in the focal spot of a 15 mW air-cooled argonion laser at 488 nm in the

sample beam of the flow cytometer. The instrument set-up was controlled daily with fluorescent reference beads (FCSC, Research Triangle Park, Raleigh-Durham, NC). Monitoring of instrument set-up for intensity and color compensation was 3-fold: (1) using lymphocytes of healthy people according to the AUTO-comp software (Becton Dickinson); (2) CD4-FITC (SK3; Becton Dickinson), CD8-PE (SKI; Becton Dickinson), CD3-PerCP (SK7; Becton Dickinson) triple staining of peripheral blood from healthy donors selected for erythrocyte or platelet transfusions; and (3) standardized fluorescent beads (Fluoro Spheres, DAKO). Fluorescence was collected at 512 to 547 nm, 572 to 591 nm, and >610 nm in the FITC, PE, and PerCP/PE-Cy5 fluorescence light channels. Fluorescence compensation was adjusted by hardware circuits. The amplification for FSC and SSC signals was linear, whereas fluorescence signals were amplified by 4-decade logarithmic amplifiers. All data were collected in list mode files and transferred to an IBM personal computer for computerized classification.

Data Analysis

The CLASSIF1 program system [6,8] (Partec, Münster, Germany) was used for list mode analysis as well as for result classification. FITC/PE, FITC/PE-Cy5, and PE/PE-Cy5 histograms were subject to quadrant analysis using fixed thresholds at one third of the 4-decade logarithmic fluorescence scales. The histograms were obtained by self-adjusting FSC/SSC gates for lymphocytes, monocytes, and granulocytes. The gates included more than 95% of all measured cells. The designation of the gate contents as lymphogate, monogate, and granulogate cells accounts for possible overlaps of, eg, lymphocytes into the FSC/SSC monogate or of monocytes into the granulogate.

Percent cell frequency, mean x, y, and the ratio y/x in each quadrant of the fluorescence histograms or within the FSC/SSC gates were evaluated for lymphogate, monogate, and granulogate cells. In addition, the relative antigen density was determined as the ratio of fluorescence and the square root of FSC as cell surface correlate. The histogram analysis resulted in 74 parameters per cell population [8], providing $3 \times 74 = 222$ parameters for lymphogate, monogate, and granulogate cells per antibody triplet. The totality of the 15 antibody triplet measurements resulted in $4 \times 222 = 888$ database columns per patient.

Data Classification

The CLASSIF1 algorithm (http://www.biochem.mpg.de/valet/classif1.html) [6,8-13] initially transforms all numbers of a given data column into –, 0, or + triple matrix characters depending on whether a number is below the lower percentile, between the lower and upper percentiles, or above the upper percentile of the value distribution in this data column for the reference patient group (healthy volunteers). Following this transformation for all data columns, a confusion matrix between the clinical diagnosis of normal, κ B-CLL, or λ B-CLL and the computer classification of normal, κ B-CLL, or λ B-CLL as calculated from the transformed data columns is established.

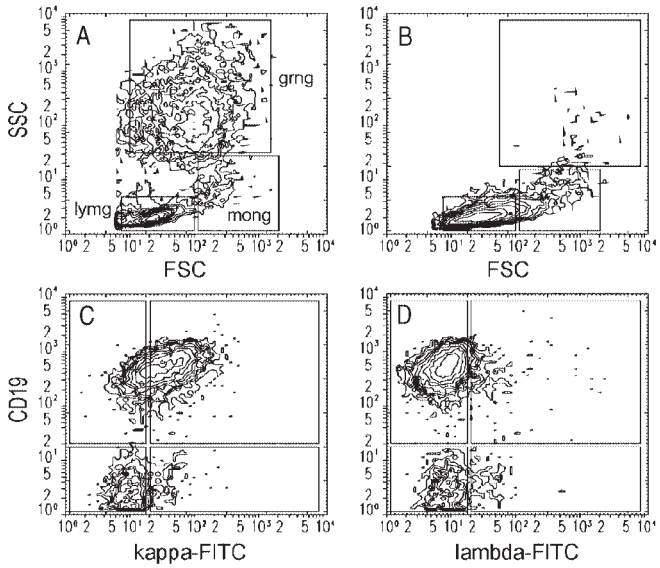

FIGURE 1. Self-adjusting forward scatter (FSC)/sideward scatter (SSC) lymphogates, monogates, and granulogates of κ (A) or λ (B)/CD19/CD5 stained normal peripheral blood leukocytes as opposed to the κ (C) and λ (D)/CD19 displays of a κ B-cell chronic lymphocytic leukemia patient being κ positive and λ negative. The autogating function is equally applicable to peripheral blood leukocytes (A) and bone marrow aspirate (B) samples. The histograms contain 9684, 9805, 7824, 7230 cells with contour lines at 10% linear steps downwards from the maximum logarithmic channel contents (178, 228, 82, 98 cells) on a 3-decade amplitude scale. The lowest contour lines contour histogram channels containing minimally 1 cell. Thus, the location of all cells in each histogram is displayed. FITC indicates fluorescein isothiocyanate.

The diagonal values of this confusion matrix represent the specificity for the correct identification of healthy reference patients and the sensitivity for the correct identification of patients with κ B-CLL or λ B-CLL. The sum of these diagonal values is maximized during the subsequent iterative learning process by exclusion of noninformative parameters, thus enriching the discriminatory parameters in the disease classification masks. The optimum classification is ideally reached when all samples are correctly classified, ie, when the value in each of the diagonal boxes of the confusion matrix is 100% and all values in nondiagonal boxes are 0%. The learning process was performed for the 10% to 90%, 15% to 85%, 20% to 80%, 25% to 75%, and 30% to 70% percentile thresholds with final selection of the optimum classification result.

The disease classification mask for the reference group of patients typically contains a sequence of (0) characters because

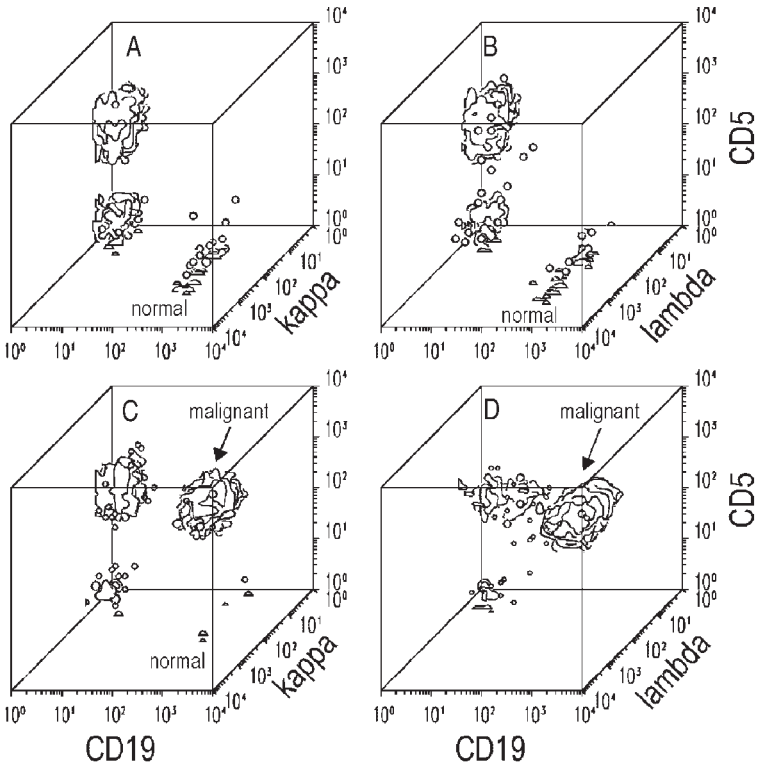

FIGURE 2. κ (A) and λ (B) staining of normal peripheral blood leukocytes (PBL) as well as κ (C) for a κ B-cell chronic lymphocytic leukemia (B-CLL) bone marrow aspirate (BMA) and λ (D) for a λ B-CLL BMA. The cell clusters represent forward scatter (FSC)/sideward scatter (SSC) lymphogate cells of κ or λ/CD19/CD5 stained samples. The FSC/SSC lymphogate location of Figures 2A and 2B is indicated in Figures 1A and 1B. The equally frequent presence of κ^+ and $\lambda^+/CD19^+/CD5^-$ cells in normal PBL (A, B) is visible. The malignant κ^+ (C) or λ^+ (D)/$CD19^+/CD5^+$ cells are well separated from a few normal $\kappa^+/CD19^+/CD5^-$ cells in the κ B-CLL (C) 3-dimensional cube. The contour lines are plotted for the 10% level of the maximum logarithmic channel content (max = 79, 19, 49, 105 cells) using a 3-decade amplitude scale. At the indicated maximum channel contents, the contour lines display the outreach of cell clusters to histogram channels with minimally 1 cell, thus assuring the display of all cells. Total contents per 3-dimensional cube were 2329, 2272, 4122, and 8659 cells.

the majority of values for each parameter of the reference group is located by definition between the 2 percentile thresholds, eg, 40% of the triple matrix characters are (0), 30% (–), and 30% (+) in the 30% to 70% percentile situation.

Unknown samples are classified according to the highest positional coincidence of the sample classification mask with any one of the previously learned disease classification masks. A patient is classified with a double classification in case of equal numbers of hits, eg, for 2 disease classification masks. Double classifications occur for transitional-state patients exhibiting borderline parameter patterns between 2 classification categories. Missing values in the triple matrix pattern or lack of overall information to clearly distinguish between the 2 classification categories are additional causes for double classifications. They represent classification errors.

All parameters contribute equally to the classification result of individual patients, and the position of the parameters in the disease classifier masks corresponds to their relative sequence in the database. The CLASSIF1 algorithm does not require assumptions on the mathematical distribution of the classified parameters; no cropping of far outreaching values is performed, and no need for the substitution of missing values exists.

RESULTS

The FSC/SSC lymphogate, monogate, and granulogate cell autogating function (Figures 1A and 1B) assures the fast evaluation of all immunophenotype histograms (Figures 1C and 1D) under standardized conditions. The simultaneous 3-color display provides a visual discrimination of particular lympho-, mono-, or granulogate cell populations, such as normal cells (Figures 2A and 2B) as well as simultaneous occurrence of normal and malignant cells (Figures 2C and 2D) in κ/CD19/CD5 or λ/CD19/CD5 assays.

κ B-CLL, λ B-CLL, and normal BMA are readily distinguishable with predictive values >95% (Table 1) by a data pattern of 19 different parameters (Table 2A). Among them,

TABLE 1. Discrimination between Normal and κ/λ B-CLL*

Clinical Diagnosis	No. of Patients	CLASSIF1 Classification, % Normal BMA	κ-CLL BMA	λ-CLL BMA
Normal BMA	22	100.0	0.0	0.0
κ-CLL BMA	26	0.0	100.0	3.8
λ-CLL BMA	27	0.0	0.0	100.0
Negative and positive predictive values, %		100.0	100.0	96.4

*Evaluation of the lymphogate cells in all antibody assays of database: C5LEARN.BI4 for the 10% to 90% percentile thresholds. B-CLL indicates B-cell chronic lymphocytic leukemia; BMA, bone marrow aspirate.

malignancy indicators are either increased or decreased in both κ and λ B-CLL, whereas malignancy discriminators differentiate between κ and λ B-CLL. The joint consideration of several parameters of the data pattern (Table 2B) as cell populations simplifies the understanding of the discriminatory changes, focusing on increased T-cells and $CD19^+/CD5^+$ lymphogate cells. Furthermore, κ^+ cells are increased in κ B-CLL and λ^+ cells in λ B-CLL.

Because BMA from healthy individuals is not always available, it was investigated whether PBL from healthy individuals could tentatively replace BMA as reference samples. The classification provides predictive values >92% (Table 3). This predictive value is similar to that of using BMA as reference samples (>95%, Table 1).

In PBL samples, even a reduced antibody panel of κ/CD19/CD5 and λ/CD5/CD5 distinguishes κ and λ B-CLL PBL from normal PBL with predictive values >95% (Table 4). Malignancy indicators and discriminators can be distinguished (Table 5A) and the parameter pattern simplified (Table 5B) as in using BMA (Table 2).

The classification permits us to assess the differences between normal BMA and PBL cells. BMA can be discriminated from PBL with predictive values >98% (Table 6). The differences concern lymphogate and granulogate cells (Table 7). Most parameters (13/15) are decreased in BMA compared to PBL except for an increase of κ on $\kappa^+/CD19^+$ and of $\%CD5^-$ cells in BMA. The differences between BMA and PBL can be used as an objective quality indicator for BMA. The differences between normal PBL and normal BMA (Table 7) do not contraindicate the use of normal PBL as reference for the κ/λ B-CLL BMA discrimination (Table 3). The observed data pattern (not shown) has similarities to the pattern of Table 2 and not to the pattern of Table 7, which would be the case if the classification in Table 3 had simply picked up the differences between normal PBL and normal BMA.

It is also possible to assess differences between κ B-CLL cells in BMA and PBL with predictive values >79% (Table 8) and differences between λ B-CLL cells in BMA and PBL with predictive values >92% (data not shown).

The clear cases among the unknown test samples were identified with predictive values >90% (Table 9) for κ/λ B-CLL in

TABLE 2. BMA: κ/λ B-CLL Disease Classifier Masks*

A. Parameter Changes

Mask Parameter No.	Parameter	Assay	Disease Classifier Masks N	κ	λ
Normal/malignant					
1	% Lymphogate cells	CD45/14/20	0	+	+
2	$\%\ CD8^-/CD4^-$	CD8/4/3	0	+	+
3	$\%\ CD4^-/CD3^-$	CD8/4/3	0	+	+
4	$\%\ CD8^-/CD3^-$	CD8/4/3	0	+	+
5	$\%\ CD19^+/CD5^+$	κ/CD19/5	0	+	+
6	$\%\ CD19^+$	λ/CD19/5	0	+	+
7	$\%\ CD8^+$	CD8/4/3	0	–	–
8	$\%\ CD8^+/CD3^+$	CD8/4/3	0	–	–
9	$\%\ CD19^-$	λ/CD19/5	0	–	–
10	$\%\ \lambda^+/CD5^-$	λ/CD19/5	0	–	–
κ/λ–CLL					
11	$\%\ \kappa^+$	κ/CD19/5	0	+	–
12	$\%\ \kappa^+/CD19^+$	κ/CD19/5	0	+	–
13	κ on $\kappa^-/CD5^+$	κ/CD19/5	0	+	0
14	$\%\ \kappa^+/CD5^+$	κ/CD19/5	0	+	0
15	$\%\ \lambda^+$	λ/CD19/5	0	–	+
16	$\%\ \lambda^+/CD19^+$	λ/CD19/5	0	–	+
17	λ on $\lambda^-/CD5^+$	λ/CD19/5	0	0	+
18	CD5/λ ratio on $\lambda^-/CD5^+$	λ/CD19/5	0	0	–
19	$\%\ \lambda^+/CD5^+$	λ/CD19/5	0	0	+

B. Cell Population Changes

Cell Population	Mask Parameter No.	Cell Population Changes
Normal/malignant		
$CD8^+/CD3^+$, $CD4^+/CD3^+$ lymphogate cells	2,3,4,7,8	Decreased
$CD19^+/CD5^+$ lymphogate cells	5,6,9	Increased
$\lambda^+/CD5^-$ lymphogate cells	10	Decreased
κ-CLL		
$\%\ \kappa^+$ lymphogate cells	11	Increased
$\%\ \kappa^+/CD19^+/CD5^+$ lymphogate cells	12,14	Increased
κ on $\kappa^+/CD5^+$ lymphogate cells	13	Increased
λ-CLL		
$\%\ \lambda^+$ lymphogate cells	15	Increased
$\%\ \lambda^+/CD19^+/CD5^+$ lymphogate cells	16,19	Increased
λ on $\lambda^-/CD5^+$ lymphogate cells	17	Increased
CD5/λ ratio on $\lambda^-/CD5^+$ lymphogate cells	18	Decreased

*The table contains the information for the lymphogate cells of the four 3-color antibody assays of database: C5LEARN.BI4. Disease classifier mask values below the 10% percentile are represented by (–), between the 10% and 90% percentile by (0), and above the 90% percentile thresholds by (+). BMA indicates bone marrow aspirate; B-CLL, B-cell chronic lymphocytic leukemia; N, normal.

TABLE 3. Discrimination between Normal PBL and κ/λ B-CLL BMA*

Clinical Diagnosis	No. of Patients	CLASSIF1 Classification, % Normal PBL	κ-CLL BMA	λ-CLL BMA
Normal PBL	58	100.0	1.7	1.7
κ-CLL BMA	26	0.0	96.2	3.8
λ-CLL BMA	27	0.0	3.7	96.3
Negative and positive predictive values, %		100.0	92.6	92.9

*Evaluation of the lympho- + mono- + granulogate cells in all antibody assays of database: F7LEARN.BI4 for the 15% to 85% percentile thresholds. PBL indicates peripheral blood leukocytes; B-CLL, B-cell chronic lymphocytic leukemia; BMA, bone marrow aspirate.

BMA (11/12) and PBL (10/10). The identification of the unknown unclear cases was consistent between BMA and PBL for patients T2, T3, and T4 (Table 10). Malignancy was recognized in patients T1 and T2, but no discrimination between κ B-CLL and λ B-CLL in the PBL samples was achieved.

DISCUSSION

The clinical utility of FCM in diagnosis of chronic lymphoproliferative disorders is well established. Accurate diagnosis of related but nevertheless distinct entities is relevant to therapeutic decisions. Correlational analysis and univariate and multivariate logistic regression models were used to determine the best combinations of antigens for the immunophenotypic classification of low-grade lymphomas [1,2]. Quantitative FCM is useful for the differential diagnosis of leukemic B-cell chronic lymphoproliferative disorders [17] but failed to show quantitative differences between atypical CLL and other lymphomas like mantle cell lymphoma [18,19].

Discrimination between normal and malignant samples is the primary goal of immunophenotype analysis in this context. The flow cytometric analysis of leukocytes in whole blood is usually performed after erythrocyte lysis and leukocyte fixation. Because lysis and fixation reagents may introduce artifacts [20], they are best avoided by the analysis of fresh whole blood samples kept at 4°C, especially during shipments. Surface immunoglobulins are the "antigens" most sensitive to undergoing storage alterations. The use of fresh samples is of particular importance for the analysis of cell membrane–bound immunoglobulins in κ or λ B-CLL.

Besides the good quality of flow cytometric specimens, the analysis critically depends on standardized staining and

TABLE 4. Discrimination between κ/λ B-CLL and Normal PBL*

Clinical Diagnosis	No. of Patients	CLASSIF1 Classification, % Normal PBL	κ-CLL PBL	λ-CLL PBL
Normal PBL	58	98.3	0.0	3.4
κ-CLL PBL	34	8.8	91.2	2.9
λ-CLL PBL	38	0.0	0.0	100.0
Negative and positive predictive values, %		95.0	96.9	95.0

*Evaluation of the lymphogate cells in the κ + λ/CD19/CD5 assay of database: CJLEARN.BI4 for the 10% to 90% percentile thresholds. Abbreviations are expanded in the footnote to Table 3.

TABLE 5. PBL: κ/λ B-CLL Disease Classifier Masks

A. Parameter Changes

Mask Parameter No.	Parameter	Assay	Disease Classifier Masks N	κ	λ
Normal/malignant					
1	% $CD19^+$	κ/CD19/5	0	+	+
2	% $CD19^+/CD5^+$	κ/CD19/5	0	+	+
3	% $\lambda^-/CD19^+$	λ/CD19/5	0	+	+
4	% $CD19^+/CD5^+$	λ/CD19/5	0	+	+
5	% $CD19^-$	κ/CD19/5	0	–	–
6	% $\kappa^-/CD19^-$	κ/CD19/5	0	–	–
7	% $CD5^-$	κ/CD19/5	0	–	–
8	% $\lambda^-/CD19^-$	λ/CD19/5	0	–	–
κ/λ-CLL					
9	κ surface density on κ^-	κ/CD19/5	0	+	0
10	% κ^+	κ/CD19/5	0	0	–
11	% $\kappa^+/CD19^+$	κ/CD19/5	0	+	0
12	λ on λ^-	λ/CD19/5	0	–	+
13	λ on $\lambda^-/CD19^+$	λ/CD19/5	0	–	+
14	% $\lambda^+/CD19^+$	λ/CD19/5	0	–	+
15	% $\lambda^+/CD5^+$	λ/CD19/5	0	0	+

B. Cell Population Changes

Cell Population	Mask Parameter No.	Cell Population Changes
Normal/malignant		
$CD19^+/CD5^+$ lymphogate cells	1-8	Increased
κ-CLL		
% κ^+ lymphogate cells	10	Unchanged
% $\kappa^+/CD19^+$ lymphogate cells	11	Increased
κ surface density on κ^- lymphogate cells	9	Increased
λ B-CLL		
% $\lambda^+/CD19^+/CD5^+$ lymphogate cells	14,15	Increased
λ on λ^- lymphogate cells	12	Increased
λ on $\lambda^-/CD19^+$ lymphogate cells	13	Increased

*The table contains the information for the lymphogate cells in the κ/CD19/CD5 and λ/CD19/CD5 assays of database: CJLEARN.BI4. Disease classifier mask values below the 10% percentile are represented by (–), between the 10% and 90% percentile by (0), and above the 90% percentile thresholds by (+). Abbreviations are expanded in the footnote to Table 3.

TABLE 6. Discrimination between Normal BMA and PBL*

Clinical Diagnosis	No. of Patients	CLASSIF1 Classification, %	
		Normal PBL	Normal BMA
Normal PBL	58	100.0	0.0
Normal BMA	22	4.5	95.5
Negative and positive predictive values, %		98.3	100.0

*Evaluation of the lymphogate + granulogate cells in all antibody assays of database: CULEARN.BI4 for the 15% to 85% percentile thresholds. Abbreviations are expanded in the footnote to Table 3.

TABLE 8. Discrimination between BMA and PBL in κ B-CLL*

Clinical Diagnosis	No. of Patients	CLASSIF1 Classification, %	
		κ-CLL PBL	κ-CLL BMA
κ-CLL PBL	34	100.0	0.0
κ-CLL BMA	26	34.6	65.4
Negative and positive predictive values, %		79.0	100.0

*Evaluation of the lympho- + mono- + granulogate cells in all antibody assays of database: D7LEARN.BI4 for the 20% to 80% percentile thresholds. Abbreviations are expanded in the footnote to Table 3.

measurement conditions and the mode of result calculation [8,9,20]. The computer-assisted manual extraction of numeric results from 2- or multidimensional cell clusters such as means, coefficients of variation, modes, or medians with standard software programs, such as Lysis II or CellQuest, is not sufficient for unambiguous classifications in the case of weak κ or λ cell membrane expression or rare malignant cells.

Progress toward generally and routinely applicable automated list mode or histogram classification programs for clinical FCM has remained limited [3-5] because of relatively complex parametric approaches. As shown by the present results, the nonparametric CLASSIF1 analysis [6] extracts the information in this study adequately from close to 900 data columns in a short time. The information content of relatively small numbers of database columns (n = 11-19; Tables 2, 5, and 7) is sufficient to distinguish unknown leukemic peripheral blood or bone marrow samples. It is also possible to differentiate patients with CLL from patients with other forms of low-grade non-Hodgkin's lymphoma [8,9] by this classification procedure.

It seems reasonable in the future to establish a series of standardized classifiers to compare automatically various leukemias and lymphomas [7-9]. The advantage of such defined classifiers is that they are principally independent of the flow cytometer and individual laboratory, provided that the flow cytometer is long-term calibrated by stable fluorescent beads (precision), the flow cytometer is capable of measuring the light scatter and fluorescence signals, and the antibodies used in the different laboratories have the same epitope specificity. The classification is furthermore stable in repetitive assays of the same cell sample or repetitive measurement of the same assay, because the coefficients of variation under these circumstances are typically on the order of 3% to 5%. The assay- and measurement-related variability is

TABLE 7. Classifier Masks for Normal BMA and PBL*

Parameter	Assay	Classifier Masks	
		PBL	BMA
Lymphogate			
% Lymphogate cells	CD8/4/3	0	–
% Lymphogate cells	CD45/14/20	0	–
CD20 on $CD20^+$	CD45/14/20	0	–
CD20 on $CD14^-/CD20^+$	CD45/14/20	0	–
κ on $\kappa^+/CD19^+$	κ/CD19/5	0	+
CD19/κ ratio on $\kappa^+/CD19^+$	κ/CD19/5	0	–
% $CD5^-$	κ/CD19/5	0	+
CD19 on $CD19^+/CD5^-$	λ/CD19/5	0	–
Granulogate			
CD19 on $CD19^+$	κ/CD19/5	0	–
CD19 on $CD19^+/CD5^+$	κ/CD19/5	0	–
CD19/κ ratio on $\kappa^+/CD19^+$	κ/CD19/5	0	–
CD5/κ ratio on $\kappa^+/CD5^+$	κ/CD19/5	0	–
CD19 on $CD19^+/CD5^+$	λ/CD19/5	0	–
CD19/λ ratio on $\lambda^+/CD19^+$	λ/CD19/5	0	–
CD5/λ ratio on $\lambda^+/CD5^+$	λ/CD19/5	0	–

*The table contains the information for the lympho- and granulogate cells of the four 3-color antibody assays of database: CULEARN.BI4. Disease classifier mask values below the 15% percentile are represented by (–), between the 15% and 85% percentile by (0), and above the 85% percentile thresholds by (+). Abbreviations are expanded in the footnote to Table 3.

TABLE 9. Unknown Test Set κ/λ B-CLL: Clear Cases

Clinical Diagnosis	No. of Patients	CLASSIF1 Classification, %	
		κ-CLL	λ-CLL
PBL cells			
κ-CLL	5	100.0	0.0
λ-CLL	5	0.0	100.0
Negative and positive predictive values, %		100.0	100.0
BMA cells			
κ-CLL	9	100.0	0.0
λ-CLL	3	33.3	100.0
Negative and positive predictive values, %		90.0	100.0

*Evaluation of the lymphogate cells in all antibody assays of databases: C5LEARN/CJLEARN.BI4 for the 10% to 90% percentile thresholds. Abbreviations are expanded in the footnote to Table 3.

TABLE 10. Unknown Test Set κ/λ B-CLL: Unclear Cases*

Patient (Samples)	CLASSIF1 Classification, %	
	BMA	PBL
T1 (0078/0719)	–	κ-CLL, λ-CLL
T2 (0884/2884)	κ-CLL	κ-CLL
T3 (0089/0009)	λ-CLL	Normal/λ-CLL
T4 (0751/0760)	κ-CLL	Normal
T5 (0737/2737)	λ-CLL	κ-CLL/λ-CLL

*Evaluation of the lymphogate cells in all antibody assays of databases: C5LEARN/CJLEARN.BI4 for the 10% to 90% percentile thresholds. Abbreviations are expanded in the footnote to Table 3.

therefore small in comparison to the variability of parameter distributions in the databases. Coefficients of variation of database parameters like % cell frequency, antigen expression, antigen expression ratios, or light scatter parameters are practically always >10%. They are frequently in the range of 30% to 70% and may be even higher in some instances.

The results inspire confidence in the practical feasibility of automated classifications in the clinical environment not only for diagnostically clear cases (Table 9) but especially for κ/λ B-CLL cases with unclear diagnosis according to conventional flow cytometric, morphological, and clinical analysis. The CLASSIF1 algorithm provides logically coherent classification results for such samples taken on different days from 2 different compartments of the same patients (Table 10) prior to treatment.

Beyond standardized diagnosis on a molecular scale, CLASSIF1 data pattern analysis offers an interesting potential for predictive medicine by multiparameter molecular analysis of cellular heterogeneity (cytomics) in organs or cellular systems (cytomes). This analysis is possible because multiparameter cytometry detects disease-induced biochemical changes at the very level of disease development and progression in the cells. At this level, data pattern classification is capable of predicting an individual patient's disease course in intensive care conditions [10,13], malignant tumors [11,12,15], and hematopoietic stem cell transplantation [14].

The approach is based on molecular cell phenotype analysis as it results from genotype and exposure. This approach seems particularly promising for repetitive controls of therapeutic success in the treatment of chronic lymphoproliferative disorders in individual patients, ie, for the dynamic prediction of therapeutic efficiency as well as for the early detection of unwanted therapeutic side effects or other complications.

REFERENCES

1. Harris NL, Jaffe ES, Stein H. A revised European-American classification of lymphoid neoplasms: a proposal from the international lymphoma study group. *Blood.* 1994;84:1361-1392.
2. Lennert K, Mohri N, Stein H, Kaiserling E. The histopathology of malignant lymphoma. *Br J Haematol.* 1975;31:193-203.
3. Thews O, Thews A, Huber Ch, Vaupel P. Computer-assisted interpretation of flow cytometry data in hematology. *Cytometry.* 1996;23:140-149.
4. Terstappen LWM, Mickaels RA, Dost R, Loken MR. Increased light scattering resolution facilitates multidimensional flow cytometric analysis. *Cytometry.* 1990;11:506-512.
5. Leary JF. Strategies for rare cell detection and isolation. *Methods Cell Biol.* 1994;42:331-358.
6. Valet G, Valet M, Tschöpe D, et al. White cell and thrombocyte disorders: standardized, self-learning flow cytometric list mode data classification with the CLASSIF1 program system. *Ann N Y Acad Sci.* 1993;677:233-251.
7. Höffkes H-G, Schmidtke G, Schmücker U, Brittinger G, Valet G. Computerized analysis of cells from patients with acute myelogenous leukemia prepared by density gradient centrifugation or erythrocyte lysis and measured by flow cytometry. *Lab Hematol.* 1995; 1:128-134.
8. Valet G, Höffkes H-G. Automated classification of patients with chronic lymphocytic leukemia and immunocytoma from flow cytometric three-color immunophenotypes. *Cytometry.* 1997;30: 275-288.
9. Bartsch R, Arland M, Lange S, Kahl C, Valet G, Höffkes H-G. Lymphoma discrimination by computerized triple matrix analysis of list mode data from three-color flow cytometric immunophenotypes of bone marrow aspirates. *Cytometry.* 2000;41:9-18.
10. Valet GK, Roth G, Kellermann W. Risk assessment for intensive care patients by automated classification of flow cytometric data. In: Robinson P, Babcock GF, eds. *Phagocyte Function: A Guide for Research and Clinical Evaluation.* New York, NY: Wiley-Liss; 1998: 289-306.
11. Van Driel BEM, Valet GK, Lyon H, Hansen U, Song JY, Van Noorden CJF. Prognostic estimation of survival of colorectal cancer patients with the quantitative histochemical assay of G6PDH activity and the multiparameter classification program CLASSIF1. *Cytometry.* 1999;38:176-183.
12. Valet G, Kahle H, Otto F, Bräutigam E, Kestens L. Prediction and precise diagnosis of diseases by data pattern analysis in multiparameter flow cytometry: melanoma, juvenile asthma and human immunodeficiency virus infection. In: Darzynkiewicz Z, Crissmann HA, Robinson JP, eds. *Cytometry.* 3rd ed. San Diego, Calif: Academic Press; 2001:487-508. *Methods in Cell Biology*; vol. 64.
13. Tarnok A, Bocsi J, Pipek M, Osmancik P, Valet G, Schneider P, Hambsch J. Preoperative prediction of postoperative edema and effusion in pediatric cardiac surgery by altered antigen expression patterns on granulocytes and monocytes. *Cytometry.* 2001;46:247-253.
14. Valet G, Cornelissen J, Lamers C, Gratama J. Predictive medicine by cytomics: outcome prediction in bone marrow stem cell transplantation (SCT). *Cytometry.* In press.
15. Valet G, MacNamara B, Mazur J, Strömberg M, Liliemark J, Porwit-MacDonald A. Individualized disease course prediction in high-grade non Hodgkin lymphoma (HG-NHL) patients. *Cytometry.* In press.

16. Jaffe ES, Harris NL, Stein H, Vardiman JW. *Tumours of Hematopoietic and Lymphoid Tissues*. Lyon, France: IARC Press; 2001:119-185.
17. Islam D, Lindberg AA, Christensson B. Peripheral blood cell preparation influences the level of expression of leukocyte cell surface markers as assessed with quantitative multicolor flow cytometry. *Cytometry*. 1995;22:128-134.
18. Veer van't MB, Kluin-Nelemans JC, Schoot van der CE, et al. Quality assessment of immunological marker analysis and the immunological diagnosis in leukemia and lymphoma: a multicenter study. *Br J Haematol*. 1992;80:458-465.
19. Brittinger G, Bartels H, Common H, et al. Clinical and prognostic relevance of the Kiel classification of non-Hodgkin's lymphomas: results of a prospective multicenter study by the Kiel Lymphoma Study group. *Hematol Oncol*. 1984;2:269-306.
20. Schmidtke G, Schmuecker U, Brittinger G, Franke A, Hoeffkes H-G. Comparative flow cytometric study of clonal excess in leukaemic peripheral blood from patients suffering from chronic lymphocytic leukaemia (B-CLL) by different antibodies, staining techniques and the effects of blood storage. *Clin Lab Haematol*. 1999;21:103-112.