

Expression and diagnostic utility of single and combined CD200, CD148 and CD160 markers in mature B-cell neoplasms as revealed by ROC and SVM analyses

AHMED F. ELSAID¹, ALAA A. OMRAN² and HEBA E. ABD ELRHMAN²

Departments of ¹Public Health and Community Medicine, and ²Clinical Pathology, Faculty of Medicine, Zagazig University, Zagazig 44519, Egypt

Received April 5, 2019; Accepted June 13, 2019

DOI: 10.3892/wasj.2019.16

Abstract. Timely and accurate diagnosis of mature B-cell neoplasm (MBN) subtypes is crucial for initiating a proper management plan. The aim of this study was to evaluate the diagnostic utility of CD200, CD148 and CD160 in different MBN subtypes using two discriminatory statistical techniques, namely receiver operating characteristic (ROC) curve analysis and support vector machine (SVM). This study included 86 patients with MBNs, 61 males and 25 females, whose data of medical histories and clinical examinations were collected. Fresh blood/bone marrow specimens were also collected and subjected to detailed morphological, cytogenetic and flow cytometric immunophenotypic examination. Diagnosis was established according to the 2017 WHO guidelines. The discriminatory performance of single and combined markers was assessed using ROC analysis. The discriminatory accuracy of the best markers was analyzed using SVM and compared to ROC analysis. The results revealed that CD148/CD200 demonstrated the best diagnostic utility in discriminating mantle cell lymphoma (MCL) from chronic lymphocytic leukemia (CLL), and hairy cell leukemia (HCL) from lymphoplasmacytic lymphoma (LPL) plus splenic marginal zone lymphoma (SMZL), with a sensitivity and specificity of 100% at signal ratios of >2.6 and <0.15, respectively. CD200 demonstrated equivalent diagnostic performance only in discriminating HCL from LPL plus SMZL at a signal cut-off value of >279. Discrimination between SMZL and LPL was unfeasible with any of the used markers. The discriminatory accuracy of SVM using CD148 and CD200 was comparable to that of ROC analysis. On the whole, the findings of this study indicate that CD148/CD200 demonstrated the best diagnostic utility in discriminating MCL from CLL,

and HCL from LPL plus SMZL. Combined markers offer a diagnostic value, particularly in difficult cases with anomalous marker expression. SVM could be used for the efficient analysis of flow cytometric data.

Introduction

Mature B-cell neoplasms (MBNs) are the most common lymphoproliferative disorders (LPDs) involving the peripheral blood (PB) and bone marrow (BM). They represent >80% of all lymphoid tumors. MBNs are a diverse group of diseases, including chronic lymphocytic leukemia (CLL), mantle cell lymphoma (MCL), follicular lymphoma (FL), splenic marginal zone lymphoma (SMZL), lymphoplasmacytic lymphoma (LPL), B-cell prolymphocytic leukemia (B-PLL) and hairy cell leukemia (HCL) (1). The accurate differential diagnosis is critical for the initiation of an effective treatment protocol.

The characterization of the cluster of differentiation (CD) expression patterns in different MBNs has permitted their utilization as molecular markers for the development of a flow cytometry-based differential diagnostic scheme for the different MBNs types. For example, the detection of CD5 surface expression in clonal B-cells has served to narrow the diagnostic possibilities among the several subtypes of B-cell lymphomas. This is due to the fact that CD5⁺ B-cells are mainly observed in CLL and MCL (2-4). Assaying for CD23 expression in addition to CD5 greatly facilitates the differential diagnosis of CLL from MCL (5), as the majority of cases of CLL are CD23⁺, whereas the majority of cases of MCL are CD23⁻.

The complete discrimination between the different subtypes of MBNs is still far from being accomplished using the currently identified CD markers. For example, the differential diagnosis of CLL from MCL could not be achieved in all cases as many CLL cases do not exhibit their known cell morphological pattern together with a weak or even no expression of CD23. Conversely, some MCL cases exhibit immunophenotypic characteristics of CLL (6). In addition, other subtypes of B-cell neoplasms may exhibit an atypical expression of CD5, thus overlapping with MCL and CLL (7-10). Further challenges arise when considering the differential diagnosis of CD5⁻ B-cell neoplasm subtypes due to the observed immunophenotypic overlap, such as that observed between the

Correspondence to: Dr Ahmed F. Elsaid, Department of Public Health and Community Medicine, Faculty of Medicine, Zagazig University, University Street, Zagazig 44519, Egypt
E-mail: afelsaid@zu.edu.eg

Key words: mature B-cell neoplasms, flow cytometry, CD200, CD148 and CD160 markers, expression pattern, discriminatory performance, support vector machine, ROC analysis

HCL and SMZL types (11,12). The imperfect discriminatory performance of the currently standard CD markers in MBNs highlights the need for use of multiple markers and/or additional markers in order to improve the overall performance of multiparameter flow cytometry-based differential diagnosis.

The CD200 marker, also known as OX-2, is a highly conserved type IA transmembrane glycoprotein that is expressed in a wide range of immunocytes, including myeloid cells, dendritic cells and B and T lymphocytes. Several studies have demonstrated that CD200 is upregulated in hematological malignancies, including multiple myeloma and acute myeloid leukemia and that it is associated with a poor prognosis (13-15). Importantly, CD200 has been shown to exhibit a differential expression pattern in different MBNs subtypes, being highly expressed in CLL and HCL, but not in MCL (16,17). This expression pattern has made CD200 a helpful marker in the differential diagnosis of the different MBN subtypes. CD148 (DEP-1 or PTPRJ) is a receptor-type protein tyrosine phosphatase that has been reported to be expressed in the majority of mature T cells and subsets of B cells. CD148 processes a growth-stimulatory activity and has been found to be differentially expressed in CD5⁺ MBN subtypes, being overexpressed in MCL, but not in CLL (18). Thus, CD148 may be a useful marker for the differential diagnosis of MCL and CLL. Another potentially useful diagnostic marker is the CD160 marker that has been reported to be highly expressed in the majority of CLL and HCL cases (19). CD160 is an Ig-like activating natural killer (NK) cell receptor that is expressed in the majority of circulating NK cells and a subset of circulating cytotoxic T cells, but is not normally expressed in B-cells (19). However, the diagnostic utility of CD160 among the different MBNs subtypes has not yet been well established.

Receiver operating characteristic (ROC) analysis has been the gold standard tool utilized to assess the discriminant performance of a diagnostic test/classifier/biomarker between 2 classes. Its algorithm depends on assessing the true-positive rate (sensitivity) and false-positive rate (1-specificity) at each biomarker threshold (20). One practical limitation of ROC analysis in hospital settings is the requirement of a single biomarker as a classifier. In the case that 2 or more biomarkers are used to enhance the accuracy of classification, methods are required to obtain a single composite parameter before applying ROC analysis. This could be achieved, for example, by using the ratio of 2 biomarkers or in the case of multiple biomarkers, by using linear discriminant analysis or logistic regression analysis (20). Support vector machine (SVM), on the other hand, utilizes a completely different algorithm to discriminate different classes by finding a decision hyperplane with maximal distance to the nearest data points (support vectors) (21). SVM can handle classification by multiple biomarkers easily by transforming data to a higher dimensional feature space using kernel function, where a discriminatory hyperplane can be found (22). Discriminatory hyperplane in p-dimensions is a p-1 dimensional subspace. In this regard, we hypothesized that SVM may be more suitable for the analysis of the complex data generated in the setting of multiparameter flow cytometry. In addition, the SVM algorithm is amenable to automation, rendering it a useful tool for the analysis of data generated in settings of multiparameter flow cytometry. However, SVM utilization in the analysis of flow cytometry data is still uncommon and its discriminatory accuracy needs to be verified against the well-established ROC analysis.

The first objective of the current study was to capitalize on the differential expression patterns of CD200, CD148 and CD160 in the different MBN subtypes to characterize their discriminatory performance, both as single markers or in combination in the multiparameter flow cytometry-based diagnostic setting. The second objective was to evaluate the performance of the SVM classifier in the differential diagnosis of MBNs against the well-established ROC analysis.

Subjects and methods

Study subjects. This study included 86 newly diagnosed patients with B-cell non-Hodgkin lymphomas, 61 males and 25 females, and ranging in age from 42 to 73 years. All patients presented to the Medical Oncology Clinic, Zagazig University Hospital, during the period from September, 2014 to June, 2017. Informed consents were obtained from all patients. The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the Institutional Review Board, Zagazig University Hospital, Faculty of Medicine. A medical history was collected from participant patients who were then subjected to detailed clinical examination. Fresh PB, and when needed BM specimens, were collected and subjected to thorough morphological, cytogenetic and flow cytometric immunophenotypic examination. Final diagnosis was established according to the 2017 WHO guidelines (1).

Inclusion and exclusion criteria. Any patients newly diagnosed B-cell non-Hodgkin lymphoma according to the WHO criteria and how had provided consent to participate in the study were included. Cases with an uncertain diagnosis, the existence of a definitive pathology, or previous chemotherapy were excluded from the study.

Flow cytometry analysis. Multicolor flow cytometric analysis was performed on fresh PB/BM specimens. All BM samples demonstrated a purity >80% with no hemodilution problems. For combined antibody (AB) staining, 100 μ l heparinized whole blood were incubated with 5 μ l of FITC-labeled mouse anti-human anti-CD19 AB (clone H1B19, cat. no. 555412, BD Biosciences) and 5 μ l of either PE-labeled anti-CD200 AB (clone MRC OX-104, cat. no. 552475, BD Biosciences), PE-labeled anti-CD148 AB (clone A3, cat. no. 328708, BioLegend), or PE-labeled anti-CD160 AB (clone BY55, cat. no. 562118, BD Biosciences) for 15 min in the dark at room temperature. Non-specific binding was determined using PE-labeled mouse anti-human IgG isotypic AB (clone G18-145, cat. no. 555787, BD Biosciences) and FITC-labeled mouse antihuman IgG₁ κ isotopic AB (Clone MOPC-21, cat. no. 555748, BD Biosciences). Following incubation, red blood cells (RBCs) were lysed with BD PharmLyse (cat. no. 555899, BD Biosciences), and lymphocytes were collected by centrifugation (250 x g for 5 min at room temperature), washed once with FACS buffer [1X phosphate buffered saline (PBS, pH 7.4), 1% bovine serum albumin (BSA), 1 mM EDTA, 0.1% sodium azide], and then re-suspended in FACS buffer.

Lymphocytes were firstly identified using forward scatter/side scatter (FSC/SSC) then gating was conducted to identify CD19⁺/CD200⁺, CD19⁺/CD160⁺, and CD19⁺/CD148⁺ cells according to the threshold established using the negative isotypic control antibodies. At least 10,000 events from each

sample were analyzed and the intensity of positive expression was evaluated using the mean fluorescence intensity (MFI). The MFI was defined as follows: A weak expression when the MFI was $<10^1$, a moderate expression when the MFI was $>10^1$ and $<10^2$, and a strong expression when the MFI was $>10^2$. Routine machine calibration and fluorescence compensation was performed on a daily basis according to the standard operating procedure of our clinical laboratory using standard fluorescence beads (CaliBRITE beads; BD Biosciences). Analysis was carried out using a BD FACSCalibur with CellQuest software (BD Biosciences) by an investigator who was blinded to the patients' clinical and laboratory data.

Statistical analysis. The distribution of MFI was examined using the Kolmogorov-Smirnov and Shapiro-Wilk tests. The results are expressed as medians $\pm 95\%$ confidence intervals. Comparisons between groups were performed using Kruskal-Wallis one-way analysis followed by the Dunn-Bonferroni post hoc test. To assess the diagnostic cut-off value for the examined markers in discriminating among the different types of MBNs, sensitivity and specificity analysis was performed using the ROC curves. To assess the diagnostic value of combining 2 markers, the MFI ratio of the 2 markers was used in ROC analysis. The area under ROC curve (AUC) was used to compare the overall discriminatory performance of different markers to assess their utility as diagnostic test. In general, an AUC >0.97 is considered excellent, between 0.93 to 0.96 is very good, between 0.75 to 0.92 is good or moderate, and <0.75 is poor and not clinically useful (23,24). Statistical comparisons of the ROC curves were made and evaluated using the method described in the study by DeLong *et al* (25).

To examine the discriminatory performance of SVM in the setting of flow cytometry-based differential diagnosis of MBN subtypes, the data were analyzed using SVM-optimized classification models and compared to ROC results. The SVM best classifier models were constructed by optimizing the C-parameter. The produced classifier model could be used for class prediction in future data.

A P-value <0.05 was considered to indicate a statistically significant difference. GraphPad Prism 6 (GraphPad Software Inc.), SPSS (Statistical package for social Scientists) version 20 (IBM Inc.) and MedCalc version 14.8.1 (MedCalc Software bvba), were used in the analysis. The freely available LIBSVM (26) functions in the e1071 package version 1.6-8 (<https://cran.r-project.org/web/packages/e1071/index.html>) were used in the open-source R statistical environment (<http://www.r-project.org>).

Results

Diagnosed MBN types. We identified 47 cases of CLL, 12 cases of MCL, 10 cases of LPL, 7 cases of SMZL, 4 cases of HCL, 3 cases of B-PLL and 3 cases of FL cases. Due to the few identified numbers, we did not include B-PLL and FL in the statistical analysis. The clinicopathological characteristics of the studied patients are presented in Table SI.

Fluorescence characteristics of the examined markers in the studied MBN cases. The distributions of the cases according to the examined markers and MFI are presented in Table I and

Fig. 1. Consistent with the findings of previous research (18), the expression of CD148 was higher in MCL vs. CLL (Fig. 1A and C). On the other hand, the expression levels of CD200 and CD160 were higher in CLL compared to MCL (Fig. 1B and C). A significantly greater percentage of CLL cases exhibited a strong CD200 and CD160 expression compared to the MCL cases (Table I). By contrast, a significantly greater percentage of MCL cases exhibited a higher CD148 expression compared to the CLL cases (Table I). As shown in Fig. 1C, the MFI was not normally distributed in all MBN cases.

SMZL exhibited a similar expression pattern to that of MCL, with a low expression of CD200 and CD160 compared to CD148 (Fig. 1C). The LPL cases exhibited a higher expression of CD148 than CD200 and CD160 (Fig. 1C). Four cases with HCL were identified in this study.

ROC analysis of the discriminatory performance of CD200, CD148 and CD160 markers in the differential diagnosis of different MBNs. ROC analysis was used to identify cut-off value, sensitivity and specificity, and the AUC of each marker diagnostic test in the differential diagnosis of MBNs (Table II and Fig. 2). CD200 yielded statistically significant fluorescent signals that could potentially aid in discriminating MCL from CLL (AUC=94%), HCL from LPL and SZML (AUC=100%), and SZML from HCL and LPL (AUC=77.6%). However, it failed to yield a statistically significant signal to discriminate LPL from HCL and SZML (AUC=56%). CD148 yielded statistically significant fluorescent signals in differentiating MCL from CLL (AUC=85.7%), HCL from LPL and SZML (AUC=97.1%), and LPL from HCL and SZML (AUC=84.5%), but failed to yield a significant signal to differentiate SZML from HCL and LPL (AUC=56%). CD160 was successful in differentiating MCL from CLL (AUC=74.7%) and HCL from LPL and SZML (AUC=83.8%).

For the comparison of the discriminatory performance of different markers, we compared the AUC of different markers (Table II). CD200 was significantly more superior than CD160 for discriminating MCL from CLL. Although CD148 demonstrated a lower performance than CD200 in discriminating MCL from CLL, it was not significantly different from CD160. For the discrimination of HCL from LPL and SZML, CD200 demonstrated the highest performance (AUC=100%) followed by CD148 (AUC=97%) and lastly by CD160 (AUC=83.8%), albeit with no significant difference between all of them (Table II). For the discrimination of LPL from HCL and SZML, CD148 and CD160 yielded moderate discriminatory performance that could negate their clinical utility (Table II). For the discrimination of SZML from HCL and SZML, CD200 was the only marker that yielded a significant signal, but with moderate discriminatory performance as judged by an AUC of 77.6%, which questions its clinical utility for this purpose.

ROC analysis of the discriminatory performance of combined markers in the differential diagnosis of different MBNs. We used the MFI ratio of any 2 markers to assess their discriminatory performance (Table III). The comparison of sensitivity, specificity and the AUC (Table III and Fig. 2) revealed that the CD148/CD200 ratio outperformed single CD200, CD148 and CD160 markers, and CD160/CD200 and CD148/CD160 ratios in differentiating MCL from CLL and HCL from LPL plus

Table I. Distribution of diagnosed mature B-cell neoplasm cases according to the level of expression of the three markers.

Marker	Expression	CLL n (%)	MCL n (%)	LPL n (%)	SMZL n (%)	HCL n (%)	FLa n (%)	B-PLL a n (%)
CD200	Strong	37 (78.7) ^b	1 (8.3)	3 (30)	2 (28.6)	4 (100)	1 (33.3)	1 (33.3)
	Weak	9 (19.9)	2 (16.7)	6 (60)	3 (42.9)	0 (0)	1 (33.3)	1 (33.3)
	Negative	1 (2.1)	9 (75.0)	1 (10)	2 (28.6)	0 (0)	1 (33.3)	1 (33.3)
CD160	Strong	27 (57.4) ^b	0 (0)	0 (0)	0 (0)	2 (50.0)	0 (0)	0 (0)
	Weak	3 (6.4)	2 (16.7)	2 (20)	1 (14.3)	1 (25.0)	1 (33.3)	0 (0)
	Negative	17 (36.2)	10 (83.3)	8 (80)	6 (85.7)	1 (25.0)	2 (50)	3 (100)
CD148	Strong	20 (42.6)	11 (91.7) ^c	9 (90)	5 (71.4)	0 (0)	1 (33.3)	2 (66.6)
	Weak	25 (53.2)	1 (8.3)	1 (10)	2 (28.6)	3 (75.0)	1 (33.3)	1 (33.3)
	Negative	2 (4.3)	0 (0)	0 (0)	0 (0)	1 (25.0)	1 (33.3)	0 (0)
Total		47	12	10	7	4	3	3

CLL, chronic lymphocytic leukemia; MCL, mantle cell lymphoma; LPL, lymphoplasmacytic lymphoma; SMZL, splenic marginal zone lymphoma; HCL, hairy cell leukemia; FL, follicular lymphoma; B-PLL, B-cell prolymphocytic leukemia. ^aFL and B-PLL were not included in further analysis due to their few numbers. ^bSignificantly higher than MCL. ^cSignificantly higher than CLL. Expression was considered strong when the MFI was >10², weak when the MFI was >10¹ and <10², and negative when the MFI was <10¹. Comparisons between groups were made using Kruskal-Wallis one-way analysis followed by the Dunn-Bonferroni post hoc test.

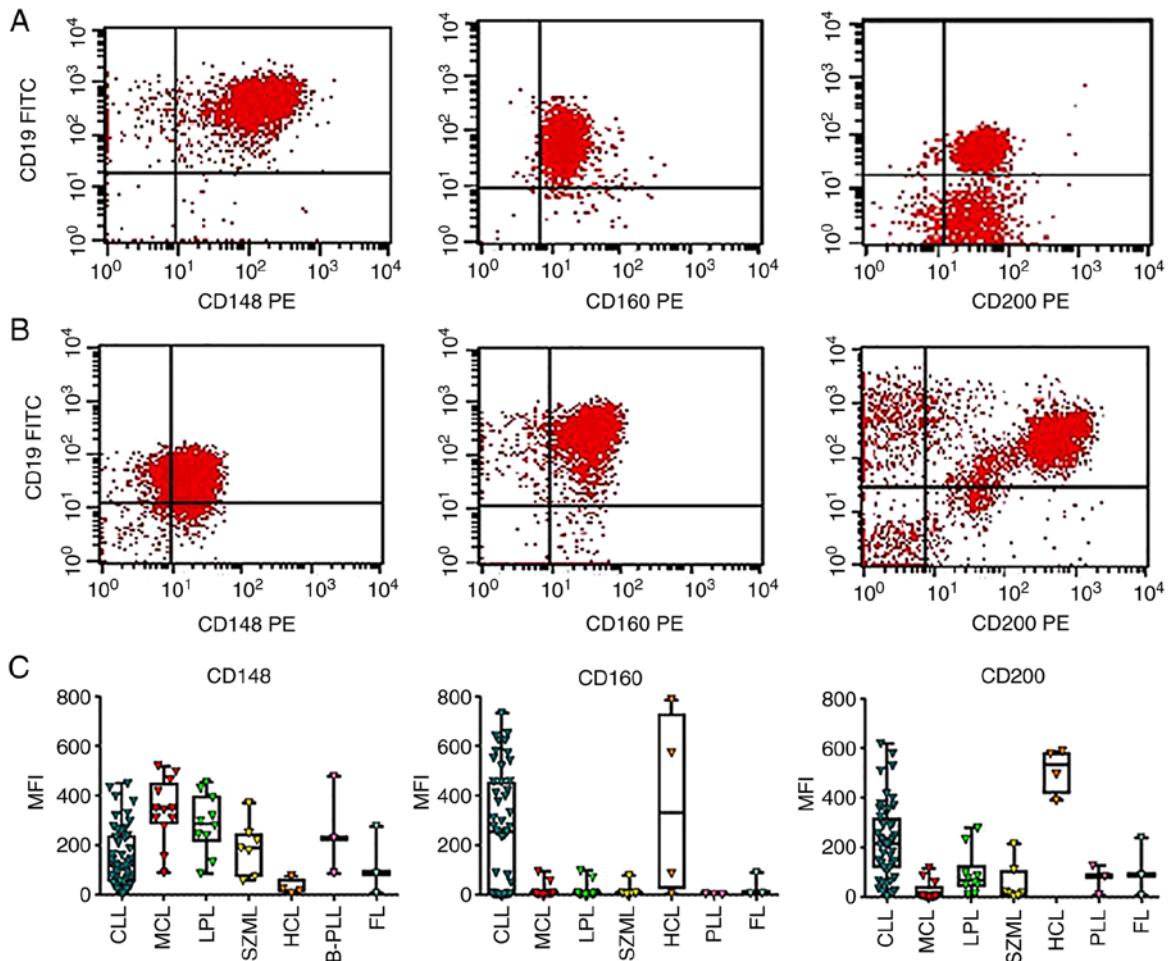


Figure 1. (A) Representative flow cytometry plot of MCL showing the expression pattern of CD148, CD160 and CD200 in CD19+ B-lymphocytes. (B) Representative flow cytometry plot of CLL showing the same 3 markers. (C) Box plot demonstrates the MFI distribution of each marker in the studied cases. MCL, mantle cell lymphoma; CLL, chronic lymphocytic leukemia.

Table II. Discriminative performance of single marker flow cytometry in differential diagnosis of different MBNs.

Marker	MCL from CLL	HCL from LPL and SMZL	LPL from HCL and SMZL	SMZL from HCL and LPL
CD200				
Cut-off	<117.5	>279	≤279	<22.8
Sensitivity	100 % (CI: 73.5-100.0)	100% (CI: 39.8-100)	100% (CI: 69.2-100)	71.43 % (CI: 29-96.3)
Specificity	74.47% (CI: 59.7-86.1)	100% (CI: 80.5-100)	36.36% (CI: 10.9-69.2)	85.71% (CI: 57.2-98.2)
AUC	94.1% ^{a,b} , P<0.001 ^d (CI: 0.847-0.985)	100% ^b , P<0.001 ^d (CI: 0.839-1.000)	56.4%, P=0.65 (CI: 0.333 to 0.776)	77.6%, P<0.05 ^d (CI: 0.543-0.926)
CD160				
Cut-off	<95.5	>78.8	≤6.2	>6.2
Sensitivity	100% (CI: 73.5-100.0)	75% (CI: 6.8-93.2)	70% (CI: 34.8-93.3)	85.7% (CI: 42.1-99.6)
Specificity	57.45 % (CI: 42.2-71.7)	94.1% (CI: 71.3-99.9)	81.8% (CI: 48.2-97.7)	57.1% (CI: 28.9-82.3)
AUC	74.7% ^c , P<0.001 ^d (CI: 0.617-0.851)	83.8% ^b , P=0.02 ^d (CI: 0.614-0.961)	73.6% ^b , P<0.05 (CI: 0.501-0.902)	53.1%, P 0.82 (CI: 0.304-0.749)
CD148				
Cut-off	>270	<77.6	> 220	≤220
Sensitivity	83.3% (CI: 51.6-97.9)	100% (CI: 39.8-100)	80% (CI: 44.4-97.5)	71.4 (CI: 29-96.3)
Specificity	83.0% (CI: 69.2-92.4)	88.24% (CI: 63.6-98.5)	81.8% (CI: 48.2-97.7)	57.1% (CI: 28.9-82.3)
AUC	85.7% ^{b,c} , P<0.001 ^d (CI: 0.742-0.935)	97.1% ^b , P<0.001 ^d (CI: 0.789-1.000)	84.5% ^b , P<0.001 ^d (CI: 0.623-0.964)	56.1%, P=0.639 (CI: 0.331-0.774)

MBNs, mature B-cell neoplasms; CLL, chronic lymphocytic leukemia; MCL, mantle cell lymphoma; LPL, lymphoplasmacytic lymphoma; SMZL, splenic marginal zone lymphoma; HCL, hairy cell leukemia; FL, follicular lymphoma; B-PLL, B-cell prolymphocytic leukemia; AUC, area under curve; CI, confidence interval. Comparisons between AUC, within each discrimination class, were made using the method of DeLong *et al* (25). ^aSignificantly higher than CD160. ^bNot significantly different from each other. ^cNot significantly different from each other. ^dSignificantly higher than a random AUC of 50%.

SMZL. Although the CD148/CD160 ratio exhibited a lower performance than the CD148/CD200 ratio, the differences were not statistically significant. In addition, CD148/CD160 was the only combination that yielded a significant AUC (82.7%) when discriminating LPL from HCL and SZML, but with moderate performance that negated its clinical utility.

Support vector machine analysis of the discriminatory performance of combined CD148 and CD200 in MBN cases. As CD148/CD200 demonstrated the best performance, we analyzed the expression pattern of the 2 markers, rather than their MFI ratio, using SVM. We aimed to find an optimal linear hyperplane (decision boundary) that could be compared with the finding of ROC analysis of the MFI ratio of the two markers. As shown in Fig. 3A, the decision boundary that best discriminated between MCL from CLL had a slope of 2.2, which agrees well with the cut-off ratio of 2.6 identified in ROC analysis. SVM discriminated MCL from CLL cases with 100% sensitivity and 97.9% specificity, which was very close to ROC analysis. As shown in Fig. 3B, SVM discriminated HCL cases from LPL and SZML with 100% sensitivity and a specificity similar to that of ROC analysis. As shown in Fig. 3B, there was difficulty in finding a decision boundary between LPL and SZML due to their overlapping expression pattern. SVM discriminated LPL from SZML and HCL with a sensitivity of 90%, but with a specificity of 57%. These results closely matched those of ROC analysis.

Discussion

Immunophenotyping by flow cytometry constitutes an important pillar of the diagnostic makeup of leukemia and lymphomas because of its simplicity, cost-effectiveness and capability of characterizing multiple cellular characteristics simultaneously. The characterization of the differential cellular expression pattern of different antigenic markers in lymphoid disorders could enhance the accuracy of flow cytometry-based diagnosis. In the current study, we evaluated the utility of using single and dual combination of CD148, CD160 and CD200 in the differential diagnosis of different types of mature B-cell neoplasms.

CD200 and CD148 exhibited reciprocal expression patterns in CLL, whereas CD160 exhibited a high expression pattern in CLL. The differential expression patterns of the 3 markers suggested that combination of 2 markers could enhance the flow cytometric-based differential diagnosis of MBNs. In this study, we first evaluated the performance of single markers in differentiating CLL from MCL. Both neoplasms are morphologically indistinguishable in many cases which complicates clinical management as these two diseases require drastically different lines of treatments. Single marker analysis confirmed the previously reported a high expression of CD200 in CLL (16,27,28). Fluorescence intensity less than the cut-off value (117.5) determined by ROC curve analysis differentiated MCL from CLL with a sensitivity of 100% and a specificity of 74.5%. CD148, in

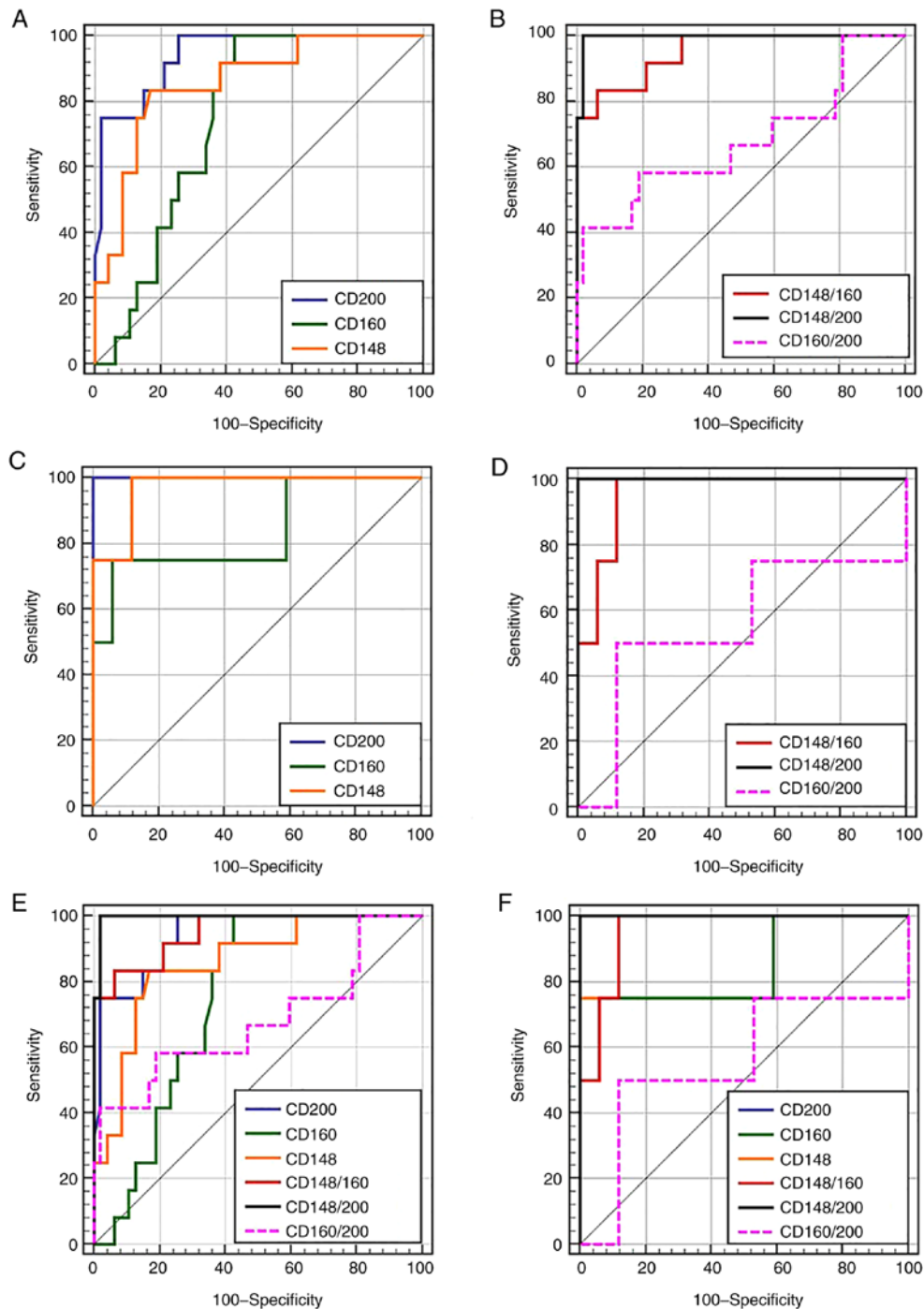


Figure 2. ROC analysis of the discriminatory performance of single and combined markers in MBN cases. (A) Discriminatory performance of single markers in the differential diagnosis of MCL from CLL. CD200 demonstrated statistically significant superior performance than CD160, but not significantly different from CD 148. (B) Discriminatory performance of combined markers in the differential diagnosis of MCL from CLL. Although CD148/CD200 demonstrated the best discriminatory performance, it was not significantly different from CD148/CD160. Both CD148/CD200 and CD148/CD160 demonstrated significantly superior performance than CD160/CD200. (C) Discriminatory performance of single marker in differential diagnosis of HCL from LPL and SZML. CD200 was not significantly different than CD148 and both were significantly superior to CD160. (D) Discriminatory performance of combined markers in the differential diagnosis of HCL from LPL and SZML. Although CD148/CD200 demonstrated the best discriminatory performance, it was not significantly different than CD148/CD160. Both were significantly superior to CD160/CD200. (E) Composite graph of (A and B) to show the comparison of all single and combined markers performance in differentiating MCL from CLL. CD148/CD200 demonstrated significantly superior performance than all single and combined markers except CD148/CD160. (F) Composite graph of (C and D) to show comparison of all single and combined markers performance in differentiating HCL from LPL and SZML. The performance of CD200 and CD148/CD200 demonstrated exactly similar performance. ROC, receiver operating characteristic; mature B-cell neoplasm; MCL, mantle cell lymphoma; CLL, chronic lymphocytic leukemia; HCL, hairy cell leukemia; LPL, lymphoplasmacytic lymphoma; SMZL, splenic marginal zone lymphoma.

contrast to CD200, exhibited a low expression pattern in MCL compared to CLL. A fluorescence intensity higher than the cut-off value (270) determined by ROC curve analysis

differentiated MCL from CLL with a sensitivity and specificity of 83.3 and 83.0%, respectively. This pattern of CD148 expression was previously reported (18,29).

Table III. Discriminative performance of combined markers in the differential diagnosis of different MBNs.

Combined marker	MCL from CLL	HCL from LPL and SMZL	LPL from HCL and SMZL	SMZL from HCL and LPL
CD148/CD200				
Cut-off	>2.63	<0.157	>0.1565	>6.6107
Sensitivity	100% (CI: 73.5-100.0)	100% (CI: 39.8-100)	100% (CI: 69.2-100)	71.43% (CI: 29-96.3)
Specificity	100% (CI: 88.7-99.9)	100% (CI: 80.5-100)	36.36% (CI: 10.9-69.2)	85.71% (CI: 57.2-98.2)
AUC	99.5% ^{a,b} , P<0.001 ^c (CI: 0.929-1.00)	100% ^{a,b} , P<0.001 ^c (CI: 0.839-1.00)	60%, P<0.466 (CI: 0.37-0.804)	73.5%, P=0.514 (CI: 0.499-0.90)
CD160/CD200				
Cut-off	>2.798	>1.1779	<0.495	>0.4206
Sensitivity	97.87% (CI: 15.2-72.3)	50% (CI: 6.8-93.2)	90% (CI: 55.5-99.7)	71.43% (CI: 29-96.3)
Specificity	41.67% (CI: 88.7-99.9)	88.24% (CI: 63.6-98.5)	54.55% (CI: 23.4-83.3)	71.43% (CI: 41.9-91.6)
AUC	67.7%, P=0.088 (CI: 0.543-0.793)	55.9%, P=0.787 (CI: 0.329-0.772)	70.9%, P=0.838 (CI: 0.473-0.88)	69.4%, P=0.1454 (CI: 0.457-0.87)
CD148/CD160				
Cut-off	>14.0	<4.777	>39.05	<39.05
Sensitivity	93.62 (CI: 82.5-98.7)	100% (CI: 39.8-100)	70% (CI: 34.8-93.3)	100 (CI: 59-100)
Specificity	83.33 (CI: 51.6-97.9)	88.24% (CI: 63.6-98.5)	100% (CI: 71.5-100)	50% (CI: 23-77)
AUC	95% ^{a,b} , P<0.001 ^c (CI: 0.860-0.990)	95.6% ^{a,b} , P<0.001 ^c (CI: 0.767-0.999)	82.7%, P<0.05 ^c (CI: 0.60-0.955)	55.1%, P=0.697 (CI: 0.32-0.765)

MBNs, mature B-cell neoplasms; CLL, chronic lymphocytic leukemia; MCL, mantle cell lymphoma; LPL, lymphoplasmacytic lymphoma; SMZL, splenic marginal zone lymphoma; HCL, hairy cell leukemia; FL, follicular lymphoma; B-PLL, B-cell prolymphocytic leukemia; AUC, area under curve; CI, confidence interval. Comparisons between AUC, within each discrimination class, were made using the method of DeLong *et al* (25). ^aSignificantly larger than CD160/CD200. ^bNot significantly different from each other. ^cSignificantly higher than a random AUC of 50%.

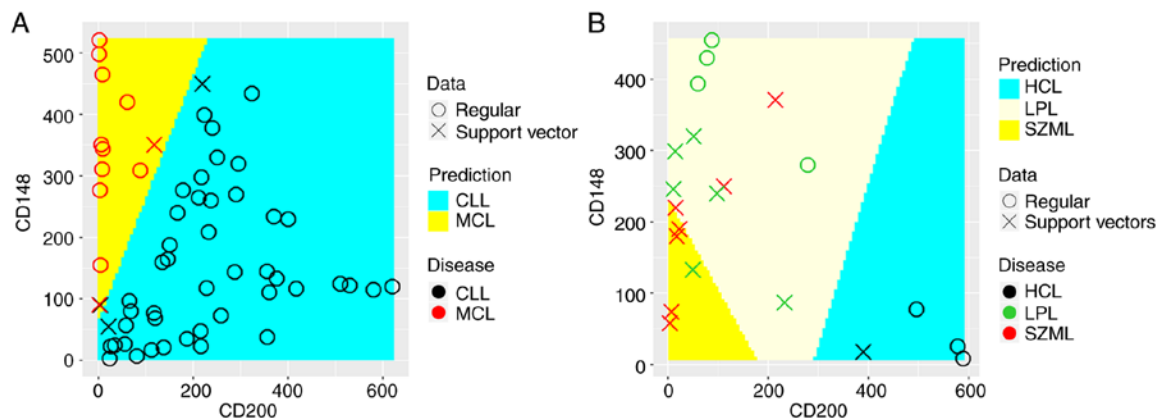


Figure 3. Support vector machine analysis of the discriminatory performance of CD148 and CD200 in MBN cases. The 'x' symbol represents data points used as support vectors to construct the discriminatory boundary between diagnosed classes, whereas the circular symbol 'o' represents the remaining data points. True classes are highlighted using symbol color (red for MCL and black for CLL or red for LPL, green for SZML and black for HCL), whereas the predicted class was highlighted using the colored background identified through the optimization process. (A) Linear SVM could accurately discriminate MCL from CLL cases with an accuracy of 98.3%. The discriminatory boundary has a slope of 2.2, which reflects the ratio of CD148/CD200. This results agrees well with that obtained by ROC analysis. (B) Multiclass classification with linear SVM using one-against-one classification strategy could accurately distinguish HCL from LPL and SZML but could not accurately distinguish LPL from SZML. This result recapitulates the results of ROC analysis. In all analyses, the tune function in package e1071 was used to optimize the cost function to provide the highest classification with least error. MBN, mature B-cell neoplasm; MCL, mantle cell lymphoma; CLL, chronic lymphocytic leukemia; HCL, hairy cell leukemia; LPL, lymphoplasmacytic lymphoma; SMZL, splenic marginal zone lymphoma.

Using the combined markers CD148/CD200 improved the distinction between MCL from CLL. A cut-off value of a ratio >2.6 was successful in discriminating MCL from CLL with 100% sensitivity and specificity. Similar results were obtained

with SVM, which utilizes a different statistical algorithm than ROC analysis. A decision boundary slope of 2.3 was found to discriminate MCL from CLL with an accuracy of 98.3% (Fig. 3A). The diagnostic utility of CD148/CD200 could be

of value to the differential diagnosis of cases of CLL with trisomy 12, in which many cases were reported to exhibit a dim CD200 expression. In such cases, the high expression of CD148, as demonstrated by the high ratio of CD148/CD200, could enhance their diagnosis. The expression pattern of CD160 was parallel to that of CD200 with a high expression in CLL and a low expression in MCL. A cut-off value of <95.5 demonstrated a perfect sensitivity of 100% in differentiating MCL from CLL, but with a very low specificity of approximately (57.5%). This result supports the finding of Lesesve *et al* (30) who recommended that CD160/200 expression levels should be included in a multiparameter approach as a second step in ambiguous cases. The combined use of CD148/CD160 significantly improved the diagnostic performance (AUC=95% vs. 74.7% in CD160 alone), but it was still lower than that of CD148/CD200 (AUC=99.5%). The combined use of CD160/CD200 demonstrated the lowest discriminative performance of MCL from CLL (AUC=67.7%; Fig. 2).

Subsequently, we evaluated the performance of each marker in the differential diagnosis of HCL from SZML and LPL. These neoplasms are difficult to be distinguished in many cases either morphologically or immunophenotypically with the standard marker panel. We observed a high CD200 MFI in HCL compared to LPL and SMZL, which was consistent with previous reports (16,17). Using ROC curve analysis of CD200 MFI, we identified a threshold value of 279 to maximally differentiate HCL from SZML and LPL. CD200 MFI >279 in our sample could perfectly discriminate HCL from SZML and LPL with 100% sensitivity and specificity. We could not establish a CD200 cut-off value to discriminate SZL and LPL. This could be explained by a similarly low CD200 expression in both SZML and LPL. Similar results were observed for CD148, in which we established a cut-off value of 77.6 to maximally differentiate HCL from LPL and SZML. CD148 MFI levels <77.6 were successful in discriminating HCL from LPL plus SZML (sensitivity 100% and specificity of 88.2%). This could be explained by the low CD148 expression in HCL, but a high expression in SZML and LPL. We could not establish a cut-off value to discriminate between LPL and SZML. The CD160 demonstrated a pattern of expression similar to CD200 with high MFI in HCL but low MFI in SZML and LPL. This pattern was previously reported by others (30,31). We established a cut-off value of 78.8 for CD160 to maximally discriminate between HCL from SZML and LPL. This resulted in moderate discriminatory performance with an AUC of 83.8%. Thus, although the CD160 expression pattern was similar to that of CD200, it demonstrated a lower performance in discriminating HCL from SZML plus LPL. The similar expression pattern in SZML and LPL did not allow establishing any cut-off value to differentiate between them.

We also assessed whether the utilization of two combined markers would improve the discrimination between HCL, LPL and SZML. ROC analysis identified a CD148/CD200 cut-off ratio of <0.15 that achieved 100% sensitivity and specificity in discriminating HCL from both LPL and SMZL. However, discrimination between LPL from SMZL was unfeasible. CD148/CD160 demonstrated a lower performance, but not with a significant difference, from CD148/CD200 in discriminating HCL from LPL and SMZL (Table III). Ratio values below the established cut-off value of approximately 4.8 demonstrated

a good discriminatory performance of HCL from LPL plus SMZL (AUC=95.6%) with a sensitivity and specificity of 100 and 88%, respectively. In contrast CD160/CD200 demonstrated poor performance in differentiating any of HCL from LPL plus SMZL or LPL from SMZL.

We evaluated the discriminatory performance of a linear SVM classifier using the CD148 and CD200 data. Given the 2-dimensional nature of our data, the discriminatory hyperplane was just a line. The SVM discriminatory accuracy of MCL from CLL cases achieved 100% sensitivity and 97.9% specificity (Fig. 3A), whereas it achieved 100% sensitivity and specificity in discriminating HCL from LPL and SZML cases (Fig. 3B). It was difficult to find a decision boundary between LPL and SZML, where the SVM discriminatory accuracy of LPL from SZML and HCL achieved a sensitivity of 90% and a low specificity of 57%. These results demonstrate that SVM classification accuracy matched that of ROC analysis.

One limitation of our study was the limited sample size. The results need to be validated in independent and larger samples. Obviously the cut-off ratios of combined markers that achieve best discriminatory performance need to be established operationally in each laboratory. In conclusion, we provided operational diagnostic criteria for the use of single or combined CD200, CD148, and CD160 markers in the differential diagnosis of different MBNs subtypes in flow cytometry setting. The combined use of CD148/CD200 has outperformed CD160/CD200, CD148/CD160, and all single markers in discriminating MCL from CLL. CD200 was as efficient as CD148/CD200 in discriminating HCL from SMZL plus LPL. Distinction between SMZL and LPL was unfeasible with the used markers and still need further research to identify suitable markers. Both SVM and ROC classifiers yielded similar discrimination accuracy in the differential diagnosis of different MBN subtypes. This result suggests that SVM may be a suitable alternative for ROC analysis given its capability of classifying multidimensional data without intermediary calculations and its amenability for automation.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

AFE contributed to the study design, flow cytometry, data analysis, table and graph presentation, the writing of the manuscript and preparing the final revision. AAO contributed to patient recruitment, sample collection, flow cytometry, and to the writing of the manuscript. HEA contributed to patient recruitment, sample collection and flow cytometry. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

Informed consents were obtained from willing participants after explaining the study purpose and design. The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by Institutional Review Board, Zagazig University Hospital, Faculty of Medicine.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J and Vardiman JW (eds): Introduction and overview of the classification of the lymphoid neoplasms. In: WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. 4th edition. IARC, Lyon, pp190-198, 2017.
- Craig FE: Flow cytometric evaluation of B-cell lymphoid neoplasms. Clin Lab Med 27: 487-512, vi, 2007.
- Craig FE and Foon KA: Flow cytometric immunophenotyping for hematologic neoplasms. Blood 111: 3941-3967, 2008.
- Stetler-Stevenson M: Flow cytometry in lymphoma diagnosis and prognosis: Useful? Best Pract Res Clin Haematol 16: 583-597, 2003.
- Matutes E, Owusu-Ankomah K, Morilla R, Garcia Marco J, Houlihan A, Que TH and Catovsky D: The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. Leukemia 8: 1640-1645, 1994.
- Kroft SH: Uncovering clinically relevant phenotypic variations in malignancies: CD23 in mantle cell lymphoma. Am J Clin Pathol 130: 159-161, 2008.
- Sánchez ML, Almeida J, Vidriales B, López-Berges MC, García-Marcos MA, Moro MJ, Corrales A, Calmuntia MJ, San Miguel JF and Orfao A: Incidence of phenotypic aberrations in a series of 467 patients with B chronic lymphoproliferative disorders: Basis for the design of specific four-color staining to be used for minimal residual disease investigation. Leukemia 16: 1460-1469, 2002.
- Baseggio L, Traverse-Glehen A, Petinataud F, Callet-Bauchu E, Berger F, Ffrench M, Couris CM, Thieblemont C, Morel D, Coiffier B, et al: CD5 expression identifies a subset of splenic marginal zone lymphomas with higher lymphocytosis: A clinico-pathological, cytogenetic and molecular study of 24 cases. Haematologica 95: 604-612, 2010.
- Asplund SL, McKenna RW, Doolittle JE and Kroft SH: CD5-positive B-cell neoplasms of indeterminate immunophenotype: A clinicopathologic analysis of 26 cases. Appl Immunohistochem Mol Morphol 13: 311-317, 2005.
- Dronca RS, Jevremovic D, Hanson CA, Rabe KG, Shanafelt TD, Morice WG, Call TG, Kay NE, Collins CS, Schwager SM, et al: CD5- positive chronic B-cell lymphoproliferative disorders: Diagnosis and prognosis of a heterogeneous disease entity. Cytometry B Clin Cytom 78 (Suppl 1): S35-S41, 2010.
- DelGiudice I, Matutes E, Morilla R, Morilla A, Owusu-Ankomah K, Rafiq F, A'Hern R, Delgado J, Bazerbashi MB and Catovsky D: The diagnostic value of CD123 in B-cell disorders with hairy or villous lymphocytes. Haematologica 89: 303-308, 2004.
- Stetler-Stevenson M and Tembhare PR: Diagnosis of hairy cell leukemia by flow cytometry. Leuk Lymphoma 52 (Suppl 2): S11-S13, 2011.
- Brunetti L, Di Noto R, Abate G, Gorrese M, Gravetti A, Raia M, Scalia G, Pascariello C, Camera A and Del Vecchio L: CD 200/OX2, a cell surface molecule with immune-regulatory function, is consistently expressed on hairy cell leukemia neoplastic cells. Br J Haematol 145: 665-678, 2009.
- Tonks A, Hills R, Whitw P, Roise B, Mils KI, Birnett AK and Darley RL: CD200 as a prognostic factor in acute myeloid leukemia. Leukemia 21: 566-568, 2007.
- Moreaux J, Hose D, Rème T, Jourdan E, Hyndemer M, Legouffe E, Moine P, Bourin P, Moos M, Corre J, et al: CD 200 is a new prognostic factor in multiple myeloma. Blood 108: 4194-4197, 2006.
- Sandes AF, de Lourdes Chauffaille M, Oliveria CR, Maekawa Y, Tamashro N, Takao TT, Ritter EC and Rizzatti EG: CD 200 has an important role in the differential diagnosis of mature B-cell neoplasms by multiparameter flow cytometry. Cytometry B Clin Cytom 86: 98-105, 2014.
- Pillai V, Pozdnyakova O, Charest K, Li B, Shahsafaei A and Dorfman DM: CD200 flow cytometric assessment and semiquantitative immunohistochemical staining distinguishes hairy cell leukemia from hairy cell leukemia-variant and other B-cell lymphoproliferative disorders. Am J Clin Pathol 140: 536-543, 2013.
- Miguet L, Béchade G, Fornecker L, Zink E, Felden C, Gervais C, Herbrecht R, Van Dorsselaer A, Mauvieux L and Sanglier-Cianferani S: Proteomic analysis of malignant B-cell derived microparticles reveals CD148 as a potentially useful antigenic biomarker for mantle cell lymphoma diagnosis. J Proteome Res 8: 3346-3354, 2009.
- Liu FT, Giustiniani J, Farren T, Jia L, Bensussan A, Gribben JG and Agrawal SG: CD160 signaling mediates PI3K-dependent survival and growth signals in chronic lymphocytic leukemia. Blood 115: 3079-3088, 2010.
- Shultz EK: Multivariate receiver-operating characteristic curve analysis: Prostate cancer screening as an example. Clin Chem 41: 1248-1255, 1995.
- Sommer C and Girlish DW: Machine learning in cell biology-teaching computers to recognize phenotypes. J Cell Sci 126: 5529-5539, 2013.
- Nayak J, Naik B and Behera HS: A comprehensive survey on support vector machine in data mining tasks: Applications & challenges. Int J Database Theory Appl 8: 169-186, 2015.
- Jones CM and Athanasiou T: Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. Ann Thorac Surg 79: 16-20, 2005.
- Fan J, Upadhye S and Worster A: Understanding receiver operating characteristic (ROC) curves. CJEM 8: 19-20, 2006.
- DeLong ER, DeLong DM and Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 44: 837-845, 1988.
- Chang CC and Lin CJ: LIBSVM: A Library for Support Vector Machines. ACM Transactions on Intelligent Systems and Technology 2: 27:1-27:27, 2011. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Palumbo GA, Parrinello N, Fargione G, Csrillo K, Chiarenza A, Berretta S, Conticello C, Villari L and Di Raimondo F: CD200 expression may help in differential diagnosis between mantle cell lymphoma and B-cell chronic lymphocytic leukemia. Leuk Res 33: 1212-1216, 2009.
- Challagundla P, Medeiros LJ, Kanagal-Shamanna R, Miranda RN and Jorgensen JL: Differential expression of CD 200 in B-cell neoplasms by flow cytometry can assist in diagnosis, sub classification and bone marrow staging. Am J Clin Pathol 142: 837-844, 2014.
- Fan L, Miao Y, Wu YJ, Wang Y, Guo R, Wang L, Shen AL, Chen YY, Xu W and Li JY: Expression patterns of CD200 and CD148 in leukemic B-cell chronic lymphoproliferative disorders and their potential value in differential diagnosis. Leuk Lymphoma 56: 3329-3335, 2015.
- Lesesve JF, Tardy S, Frotscher B, Latger-Cannard V, Feugier P and De Carvalho Bittencourt M: Combination of CD160 and CD200 as a useful tool for differential diagnosis between chronic lymphocytic leukemia and other mature B-cell neoplasms. Int J Lab Hematol 37: 486-494, 2015.
- Farren TW, Giustiniani J, Liu FT, Tsitsikas DA, Macey MG, Cavenagh JD, Oakervee HE, Taussig D, Newland AC, Calaminici M, et al: Differential and tumor-specific expression of CD160 in B-cell malignancies. Blood 118: 2174-2183, 2011.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.