Combining Machine Learning (ML) with Flow Cytometry Based Immunophenotyping: a Novel Liquid Biopsy Aid for Identifying the Risk for Prostate Cancer (PCa)

George A. Dominguez[1*], Alexander Polo[1], John Roop[1], Anthony Campisi[1], Dmitry I. Gabrilovich[2], Amit Kumar[1]

[1]Anixa Biosciences, Inc., San Jose, CA; [2]The Wistar Institute, Philadelphia, PA

BACKGROUND AND PURPOSE

- Blood-based biomarkers that can accurately predict prostate cancer (PCa) in at-risk men are lacking and result in a large percentage of men undergoing unnecessary prostate biopsy procedures (false positives).
- Myeloid-derived suppressor cells (MDSCs) are known to be key contributors in supporting tumor progression and tumor escape with several studies quantifying them in order to detect tumor development, monitor progression, and/or predict therapeutic responses.
- Here, we have developed a novel methodology for analyzing flow cytometry data using machine learning (ML) with pattern recognition neural networks (PRNN) to predict whether an at-risk male is at greater risk for PCa (Gleason Score 7 or higher) based upon the immunophenotyping of MDSCs and various other myeloid and lymphocyte populations.

METHODS

We used standard multiparametric flow cytometry techniques to immunophenotype MDSC, myeloid, and lymphocyte cell populations (**Figure 1**) found in the peripheral blood of 151 biopsy-verified PCa (Gleason Score 6 = 57; >GS 6 = 94) and 271 biopsy-verified benign prostatic hyperplasia (BPH) subjects. Peripheral blood mononuclear cells (PBMCs) were isolated 20 to 36 hours post-collection using standard methods. Myeloid and lymphocyte cell populations were analyzed on a BD LSRII flow cytometer. Compensated channel values for each event were exported from FCS format to CSV format using FlowJo software.

To prepare the data for input to the PRNN, we used a procedure wherein each channel from the FCS data export file for a panel was used as an axis in a multidimensional space. Each axis was then divided into four segments and each event was defined by its segment location within each axis. The number of total segments is a function of the number of channels used. For example, if the flow cytometry data has 4 channels, each channel is divided into four segments thus resulting in 4^4 or 256 discrete regions in the multidimensional space; these regions are referred to as hypervoxels (illustrated in **Figure 2**).

For our study, two separate panels were used with one containing seven channels (lymphoid) and one containing nine channels (myeloid) resulting in the panels having 4^7, or 16,384, hypervoxels and 4^9, or 262,144, hypervoxels, respectively. A count was then made of the number of events falling within each hypervoxel resulting in a common feature for all samples. Each hyperspace was then converted to a one-dimensional vector and used as input for a separate PRNN (**Figure 3**). After data preparation, a series of PRNNs were created with network inputs consisting of the numerical event counts in each hypervoxel; three datasets were then constructed: the training dataset – to 'teach' the two output categories through backpropagation and parameter fitting; the validation dataset – to evaluate the fit to minimize overfitting; and the test dataset – to rank the trained networks against each other and estimate the classification performance.

Figure 1. **Manual gating analysis of various myeloid and lymphocyte populations.**

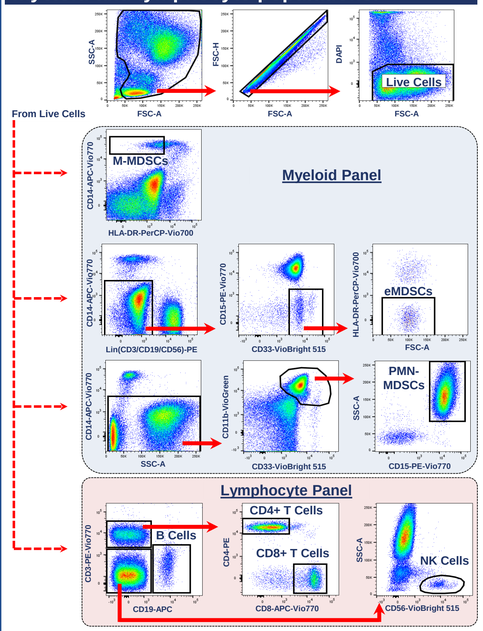


Figure 2. **Illustration of the creation of hypervoxels.**

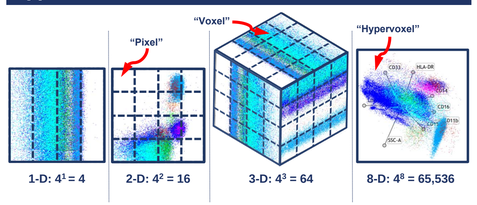
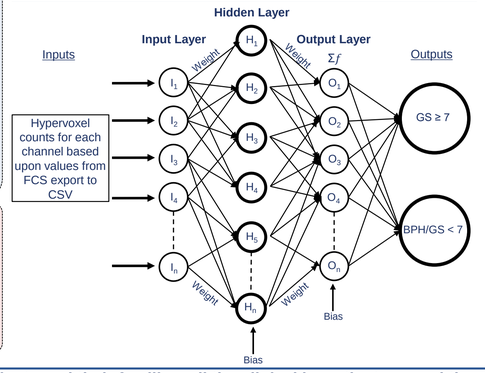


Figure 3. **Illustration of a pattern recognition neural network (PRNN).**



Finally, a holdout testing set (i.e. never seen by the network) was used to determine the overall performance of a voting ensemble of the top-ranking networks. PRNNs were trained with raw flow cytometry data from two data sets: subjects with PCa GS≥7 vs subjects with BPH/PCa GS<7. Predictions were evaluated using the performance of the trained PRNN ensemble on 51 known PCa GS≥7 and 66 known BPH/PCa GS<7 not used in the PRNN training set.

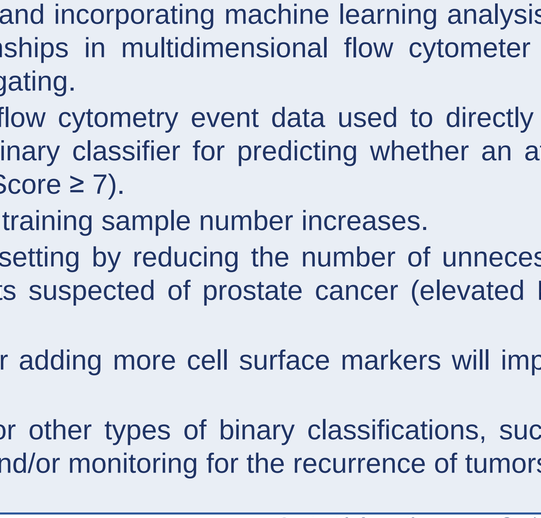
ROC curve and subsequent sensitivity and specificity analysis was performed using GraphPad Prism. This was not a blinded study and was IRB approved at each participating site.

RESULTS

Table 1. Sensitivity and specificity analysis of the PRNN for PCa classification

		Holdout		All	
		GS ≥7	BPH/ GS6	GS ≥7	BPH/ GS6
		Actual			
		51	66	94	177
Predicted	GS ≥7	47	39	87	106
	BPH/GS6	4	27	7	71
Sensitivity (95%CI)		*92.16% (81.12 to 97.82)*		92.55% (85.26 to 96.95)	
Specificity (95%CI)		*40.91% (28.95 to 53.71)*		40.11% (32.82 to 47.73)	
PPV (95%CI)		*54.65% (49.26 to 59.93)*		45.08% (41.80 to 48.40)	
NPV (95%CI)		*87.10% (71.61 to 94.75)*		91.03% (82.94 to 95.49)	
Accuracy (95%CI)		*63.25% (53.84 to 71.97)*		58.30% (52.18 to 64.24)	

Figure 4. **Receiver operating characteristic plot of the PRNN analysis for PCa classification.**



- ✓Using a PRNN, we were able to predict *47 out of 51 samples as GS≥7* and *27 out of 66 samples as BPH/GS6* for a sensitivity and specificity of 92.16% (95%CI 81.12% to 97.82%) and 40.91% (95%CI 28.95% to 53.71%), respectively (**Table 1**).
- ✓ROC curve analysis resulted in an AUC of 0.6788 (95%CI 0.5811 to 0.7766) for Holdout Samples and 0.7461 (95%CI 0.6857 to 0.8064) for All Samples (**Figure 4**).

CONCLUSIONS

- By using hypervoxel counts as a common feature and incorporating machine learning analysis, we are able to evaluate numerous complex relationships in multidimensional flow cytometer data simultaneously that cannot be done using manual gating.
- To our knowledge, this is the first application of flow cytometry event data used to directly train pattern recognition neural networks to create a binary classifier for predicting whether an at-risk male has a high risk for prostate cancer (Gleason Score ≥ 7).
- We expect network performance to improve as the training sample number increases.
- This technique may prove beneficial in a clinical setting by reducing the number of unnecessary prostate biopsies performed each year on patients suspected of prostate cancer (elevated PSA, abnormal DRE, etc.).
- We are expanding our study to determine whether adding more cell surface markers will improve the performance.
- We believe that this technology could be used for other types of binary classifications, such as predicting patient responses to immunotherapies and/or monitoring for the recurrence of tumors.

Acknowledgements: The authors would like to thank the patients and their families, all the clinical investigators, and the study staff from each participating sites.

Contact information: george@anixa.com

Presented at the 34th Congress of the International Society for Advancement of Cytometry; June 22-26, 2019; Vancouver, British Columbia, Canada