

Using Artificial Intelligence to Distinguish Subjects with Prostate Cancer (PCa) from Benign Prostatic Hyperplasia (BPH) Through Immunophenotyping of MDSCs and Lymphocyte Cell Populations

George A. Dominguez^{1*}, John Roop¹, Alexander Polo¹, Anthony Campisi¹, Dmitry I. Gabrilovich², Amit Kumar¹

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

BACKGROUND AND PURPOSE

- ❖ The purpose of this study was to determine if supervised machine learning (ML) could be used to classify samples using the immunophenotyping flow cytometry data of myeloid-derived suppressor cells (MDSCs) and lymphoid cells from healthy donor (HD), benign prostatic hyperplasia (BPH)/lower risk prostate cancer (LR-PCa), and higher risk prostate cancer (HR-PCa) subjects.
- ❖ MDSCs are a heterogeneous group of immature myeloid cells that are known to elicit immunosuppressive behavior
- ❖ The levels of MDSCs found in blood have been shown to increase over time with tumor progression and could be possible predictors for immunotherapy responses.
- ❖ Here, we created an analytical technique for processing flow cytometry data to serve as inputs for feedforward neural networks (NNs) to classify samples based upon their levels of MDSCs and lymphocytes to classify them as either HD, BPH, or PCa.

METHODS

With IRB approval, peripheral blood was collected from 203 subjects prior to a transrectal ultrasound guided prostate (TRUSP) biopsy and 116 tumor free control donors (HD). Subjects were excluded from the study if they had a previous history of cancer (excluding active surveillance), had a medical intervention for prostate cancer, or were receiving a dihydrotestosterone DHT/alpha-1 blocker for active treatment of BPH. A total of 114 subjects with biopsy confirmed prostate cancer (PCa) and 89 subjects with biopsy confirmed BPH were included (**Table 1**). Peripheral blood mononuclear cells (PBMCs) were isolated 20 to 30 hours post-collection using standard methods (**Figure 1a**). Myeloid and lymphocyte cell populations were analyzed on a BD LSRII flow cytometer (**Table 2**). Compensated channel values for each event were exported from FCS format to CSV format using FlowJo software.

For our study, two separate panels were used with one containing seven channels (lymphoid) and one containing ten channels (myeloid) (**Figure 1b**). To prepare the data for neural network (NN) input, we used a procedure wherein each channel of flow cytometry data was used as an axis in a multidimensional space (**Figure 1c**). The neural network (NN) is *trained* by inputting known classes of samples and repeatedly having its determination of the sample classification corrected. Numerous neural networks were created with each undergoing their own training cycles using three datasets: the training dataset 'teaches' the two output categories through backpropagation and fitting; the validation dataset evaluates the fit to minimize overfitting; and the test dataset ranks the trained networks against each other and estimates performance.

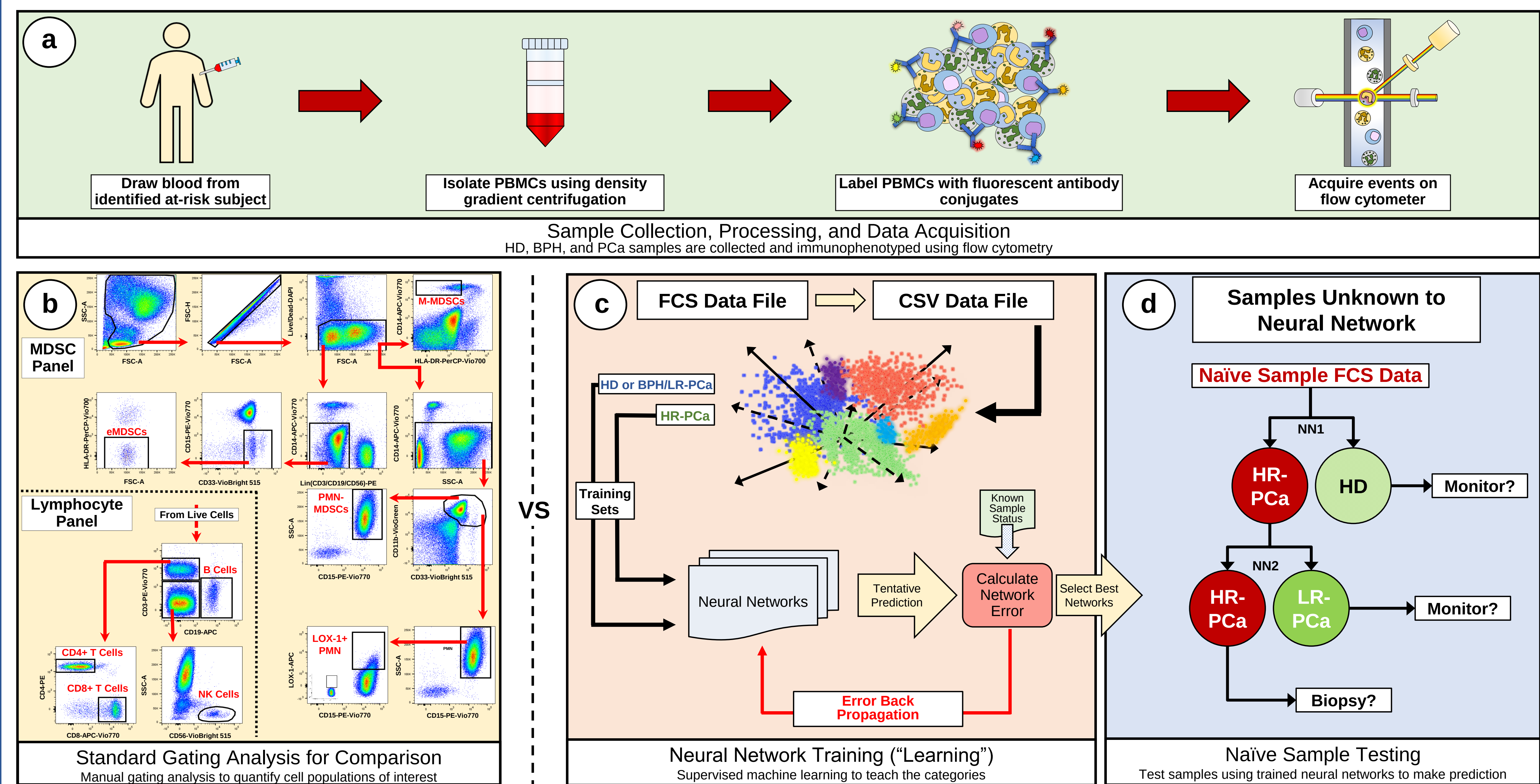
The inputs are the numerical event counts from each channel (fluorescent and SSC-A) with the output being the final classification. For sensitivity analysis, a set of test samples previously withheld from the NN training process were evaluated and statistics determined.

Additionally, a two-network approach was developed to predict whether a subject was at higher risk for PCa (HR-PCa) and should be recommended for biopsy ($G \geq 4+3$) or be actively monitored (BPH or $G \leq 3+4$). NN1 predicts whether the sample looks like HD or HR-PCa. If the sample is predicted to be a possible HR-PCa, then it is tested by NN2 and predicted as either benign/at lower risk (Benign/LR-PCa) for PCa or HR-PCa (**Figure 1d**). The final decision for biopsy would be made by the clinician.

METHODS

Continued

Figure 1. Illustrated overview of methods in this study from sample collection and analysis, neural network creation and training, and final classification.



RESULTS

Table 1. Clinical characteristics of HD, BPH, and PCa subjects

	HD	BPH	PCa
Number	116	89	114
Median Age	52	62	67
Age Range	18-79	40-81	42-86
Gleason 6			44
Gleason 7(3+4)			26
Gleason 7(4+3)			22
Gleason 8 or higher			22
Stage T1c			75
Stage T2a			5
Stage T2c			2
Stage Unknown			32

✓ Manual gating of MDSCs show trends with substantial overlap and are not specific enough to make any confident predictions (**Figure 2**).

✓ ROC analysis shows that neural network testing (AUROC = 0.929) is far improved over manual gating of MDSCs to distinguish between HD and HR-PCa (**Figure 3**)

✓ Sensitivity and specificity of 90% and 96%, respectively, were achieved to distinguish naïve HD and HR-PCa samples (**Table 3**). For all samples, a sensitivity and specificity of 84% and 95%, respectively, were achieved.

✓ For samples that are either biopsy confirmed BPH or PCa ($G \geq 6$), the networks identified 9 out of 10 naïve samples correctly as needing a biopsy (*Biopsy*; $G > 3+4$) and 26 out of 50 naïve samples correctly as not needing a biopsy (*No Biopsy*; BPH/LR-PCa) resulting in a sensitivity and specificity of 90% and 52%, respectively (**Table 4**).

✓ For all subjects who underwent a TRUSP biopsy, 37 out of 44 were correctly recommended as *Biopsy* and 108 out of 159 samples were correctly recommended as *No Biopsy*.

Figure 2. Manual gating analysis of the three primary subsets of MDSCs: eMDSCs (left graph), PMN-MDSCs (middle graph), and M-MDSCs (right graph).

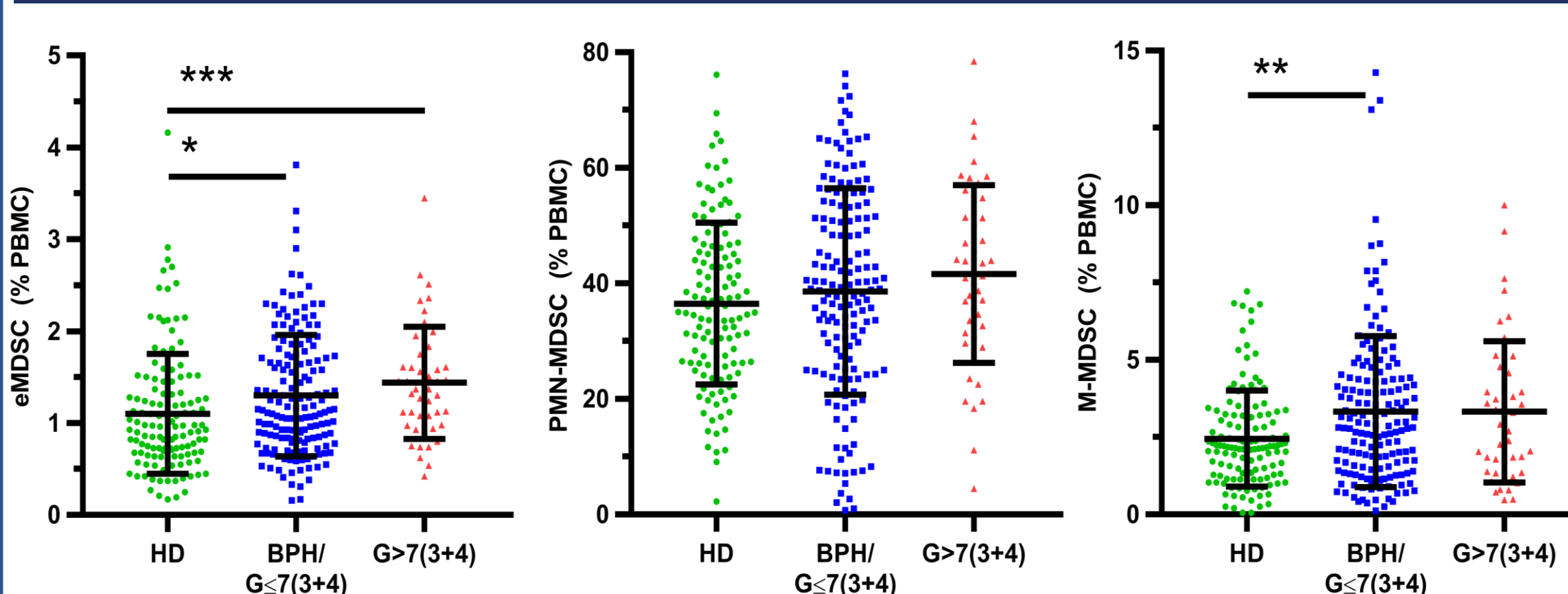


Table 3. Sensitivity/specificity analysis for predicting HD vs HR-PCa (NN1)

	Naïve		All	
	HR-PCa	HD	HR-PCa	HD
Samples Tested	10	50	44	116
Predicted	HR-PCa	9	2	37
	HD	1	48	7
Sensitivity	90%		84%	
Specificity	96%		95%	
Accuracy	95%		92%	

Table 4. Sens. and Spec. two-network test to distinguish between BPH/LR-PCa and HR-PCa

	Naïve		All	
	HR-PCa	BPH/LR-PCa	HR-PCa	BPH/LR-PCa
Samples Tested	10	50	44	159
Recommended	Biopsy	9	24	37
	No Biopsy (Monitor)	1	26	7
Sensitivity	90%		84%	
Specificity	52%		68%	
Accuracy	58%		71%	

CONCLUSIONS

- To our knowledge, this is the first application for flow cytometry event data used to directly train artificial neural networks to detect higher risk prostate cancer and classify subjects based upon the immunophenotyping of MDSCs and lymphocytes.
- We expect classification to improve as more samples are included in the training sets.
- With over 1 million prostate biopsies performed each year, we believe this technique could be used to dramatically reduce the number of unnecessary prostate biopsies performed each year on high-risk patients (high PSA, abnormal DRE, age) as current screening methods result in a large number of false positives.
- Furthermore, 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.

Figure 3. ROC curve comparing manual gating to neural network for HD vs HR-PCa naïve test samples

