

Table 1S. Analysed variables and corresponding thresholds in – off-therapy subjects. Set = training or test sets; AUC = Area Under the Receiver Operating Characteristics curve; MCC = Matthew’s Correlation Coefficient; Sens = sensitivity; Spec = specificity; PPV = positive predictive value; NPV = negative predictive value; F = F-measure; CA = Classification Accuracy.

Variable	Threshold	Set	AUC	MCC	Sens	Spec	PPV	NPV	F
CD4CXCR3BAL/PB	≤ 1.43	Training Set	0.91	+0.76	0.87	0.90	0.81	0.93	0.84
		Test Set	0.96	+0.78	0.86	0.92	0.86	0.92	0.86
CD4CXCR3%BAL	≤ 60.5	Training Set	0.88	+0.70	0.80	0.90	0.80	0.90	0.80
		Test Set	0.88	+0.39	0.71	0.69	0.56	0.82	0.63
CD4CXCR3/CCR4BAL	≤ 1.39	Training Set	0.98	+0.69	0.67	0.97	0.91	0.85	0.77
		Test Set	0.96	+0.30	0.43	0.85	0.60	0.73	0.50
CD4CCR4BAL/PB	> 1.2	Training Set	0.88	+0.66	0.80	0.87	0.75	0.90	0.77
		Test Set	0.53	-0.03	0.43	0.54	0.33	0.64	0.38
Neutrophils (%) BAL	> 6.2	Training Set	0.85	+0.64	0.73	0.90	0.79	0.87	0.76
		Test Set	0.88	+0.60	0.86	0.77	0.67	0.91	0.75
CD4CCR4 (%) BAL	> 34.7	Training Set	0.85	+0.62	0.80	0.83	0.71	0.89	0.75
		Test Set	0.60	+0.04	0.43	0.62	0.38	0.67	0.40
Lymphocytes (%) BAL	≤ 6.9	Training Set	0.82	+0.57	0.93	0.67	0.58	0.95	0.72
		Test Set	0.79	+0.60	1.00	0.62	0.58	1.00	0.74
Eosinophils (%) BAL	> 3.3	Training Set	0.75	+0.54	0.80	0.77	0.63	0.88	0.71
		Test Set	0.70	+0.43	0.57	0.85	0.67	0.79	0.62
CD8CXCR3BAL/PB	≤ 1.42	Training Set	0.63	+0.38	0.67	0.73	0.56	0.81	0.61
		Test Set	0.95	+0.48	1.00	0.46	0.50	1.00	0.67
CD4/CD8BAL	≤ 1.6	Training Set	0.68	+0.38	0.73	0.67	0.52	0.83	0.61
		Test Set	0.46	+0.03	0.57	0.46	0.36	0.67	0.44

CD8CXCR3 (%) BAL	≤ 89.7	Training Set	0.61	+0.24	0.93	0.27	0.39	0.89	0.55
		Test Set	0.76	+0.37	1.00	0.31	0.44	1.00	0.61
CD8CCR4 (%) BAL	> 3.5	Training Set	0.44	+0.24	0.93	0.27	0.39	0.89	0.55
		Test Set	0.63	-0.06	0.71	0.23	0.33	0.60	0.45
CD8CCR4BAL/PB	> 0.38	Training Set	0.56	+0.24	0.93	0.27	0.39	0.89	0.55
		Test Set	0.62	+0.10	0.86	0.23	0.38	0.75	0.52
CD8CXCR3/CCR4BAL	≤ 16.6	Training Set	0.45	+0.21	0.87	0.33	0.39	0.83	0.54
		Test Set	0.68	+0.31	1.00	0.23	0.41	1.00	0.58
Macrophages (%) BAL	> 64.5	Training Set	0.51	+0.20	0.80	0.40	0.40	0.80	0.53
		Test Set	0.70	-0.12	0.57	0.31	0.31	0.57	0.40

Table 2S: Variables and corresponding threshold analysed in off therapy subjects (excluding SAR subjects). Set = training or test sets; AUC = Area Under the Receiver Operating Characteristics curve; MCC = Matthew's Correlation Coefficient; Sens = sensitivity; Spec = specificity; PPV = positive predictive value; NPV = negative predictive value; F = F-measure; CA = Classification Accuracy.

Variable	Threshold	Set	AUC	MCC	Sens	Spec	PPV	NPV	F
CD4CXCR3BAL/PB	≤ 1.43	Training Set	0.91	0.70	0.87	0.83	0.81	0.88	0.84
		Test Set	0.96	0.73	0.86	0.88	0.86	0.88	0.86
CD4CXCR3/CCR4BAL	≤ 1.39	Training Set	0.89	0.65	0.67	0.94	0.91	0.77	0.77
		Test Set	0.76	0.19	0.43	0.75	0.60	0.60	0.50
CD4CXCR3 (%) BAL	≤ 60.5	Training Set	0.88	0.63	0.80	0.83	0.80	0.83	0.80
		Test Set	0.88	0.46	0.71	0.75	0.71	0.75	0.71
CD4CCR4BAL/PB	> 1.2	Training Set	0.88	0.63	0.80	0.83	0.80	0.83	0.80
		Test Set	0.53	-0.33	0.43	0.25	0.33	0.33	0.38
CD4CCR4 (%) BAL	> 31	Training Set	0.85	0.59	0.87	0.72	0.72	0.87	0.79
		Test Set	0.60	-0.19	0.57	0.25	0.40	0.40	0.47
Neutrophils (%) BAL	> 6.2	Training Set	0.85	0.57	0.73	0.83	0.79	0.79	0.76
		Test Set	0.88	0.61	0.86	0.75	0.75	0.86	0.80
Lymphocytes (%) BAL	≤ 6.9	Training Set	0.82	0.56	0.93	0.61	0.67	0.92	0.78
		Test Set	0.79	0.66	1.00	0.63	0.70	1.00	0.82
Eosinophils (%) BAL	> 3.3	Training Set	0.75	0.52	0.80	0.72	0.71	0.81	0.75
		Test Set	0.70	0.33	0.57	0.75	0.67	0.67	0.62
CD8CCR4BAL/PB	> 0.38	Training Set	0.56	0.37	0.93	0.39	0.56	0.88	0.70
		Test Set	0.62	0.13	0.86	0.25	0.50	0.67	0.63
CD4/CD8BAL	≤ 1.6	Training Set	0.68	0.29	0.73	0.56	0.58	0.71	0.65

			Test Set	0.46	-0.19	0.57	0.25	0.40	0.40	0.47
CD8CCR4 (%) BAL	>	3.9	Training Set	0.44	0.29	0.87	0.39	0.54	0.78	0.67
			Test Set	0.63	-0.04	0.71	0.25	0.45	0.50	0.56
CD8CXCR3BAL/PB	≤	1.42	Training Set	0.63	0.28	0.67	0.61	0.59	0.69	0.62
			Test Set	0.95	0.47	1.00	0.38	0.58	1.00	0.74
Macrophages (%) BAL	>	63.5	Training Set	0.54	0.26	0.80	0.44	0.55	0.73	0.65
			Test Set	0.64	-0.05	0.57	0.38	0.44	0.50	0.50
CD8CXCR3/CCR4BAL	≤	16.6	Training Set	0.45	0.23	0.87	0.33	0.52	0.75	0.65
			Test Set	0.68	0.37	1.00	0.25	0.54	1.00	0.70
CD8CXCR3 (%) BAL	≤	89.7	Training Set	0.61	0.22	0.93	0.22	0.50	0.80	0.65
			Test Set	0.76	0.47	1.00	0.38	0.58	1.00	0.74

Table 3S: Comparison between the discriminative performances of the best discretized variables and multivariate classification models in off-therapy subjects. Algorithm = single variable or model (CT = Classification Trees; CIT = Conditional Inference Trees; RFs = Random Forests; NB = Naïve Bayes; LR = Logistic Regression; LASSO = Least Absolute Selection and Shrinkage Operator)

Identified models: a CD4CXCR3BAL/PB <1.14; b CD4CXCR3%BAL, Smoking history; c Macrophages (%) BAL + Lymphocytes (%) BAL + Neutrophils (%) BAL + Eosinophils (%) BAL + CD4CCR4 (%) BAL + CD8CCR4 (%) BAL + CD4CCR4BAL/PB + sex; d Lymphocytes (%) BAL + CD4CXCR3 (%) BAL + CD4CCR4 (%) BAL + CD4CCR4BAL/PB; * all variables were included.

Algorithm	Cohort	MCC	Sens	Spec	PPV	NPV	F
CD4CXCR3BAL/PB ≤1.43	Training Set	0.76	0.87	0.90	0.81	0.93	0.84
	Test Set	0.78	0.86	0.92	0.86	0.92	0.86
CT ^a	Training Set	0.76	0.67	1.00	1.00	0.86	0.80
	Test Set	0.42	0.43	0.92	0.75	0.75	0.55
CIT ^b	Training Set	0.70	0.80	0.90	0.80	0.90	0.80
	Test Set	0.39	0.71	0.69	0.56	0.82	0.63
RF [*]	Training Set	1.00	1.00	1.00	1.00	1.00	1.00
	Test Set	0.52	0.86	0.69	0.60	0.90	0.71
NB [*]	Training Set	0.71	0.87	0.87	0.76	0.93	0.81
	Test Set	0.56	0.71	0.85	0.71	0.85	0.71
LR [*]	Training Set	1.00	1.00	1.00	1.00	1.00	1.00
	Test Set	0.39	0.71	0.69	0.56	0.82	0.63
LR (stepwise) ^c	Training Set	1.00	1.00	1.00	1.00	1.00	1.00
	Test Set	0.30	0.43	0.85	0.60	0.73	0.50
LASSO ^d	Training Set	0.71	0.87	0.87	0.76	0.93	0.81
	Test Set	0.39	0.86	0.54	0.50	0.88	0.63