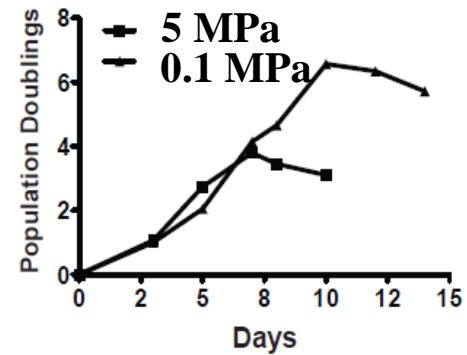
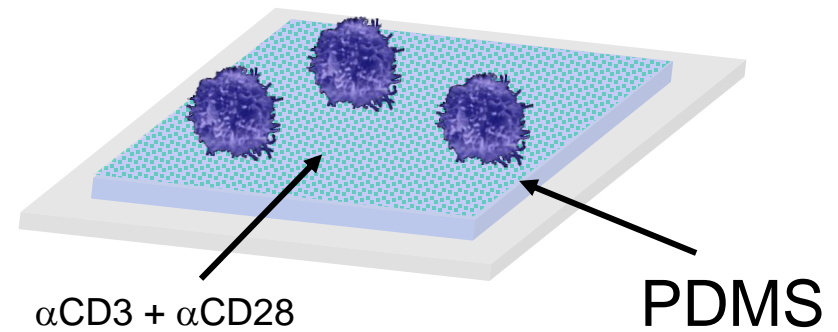
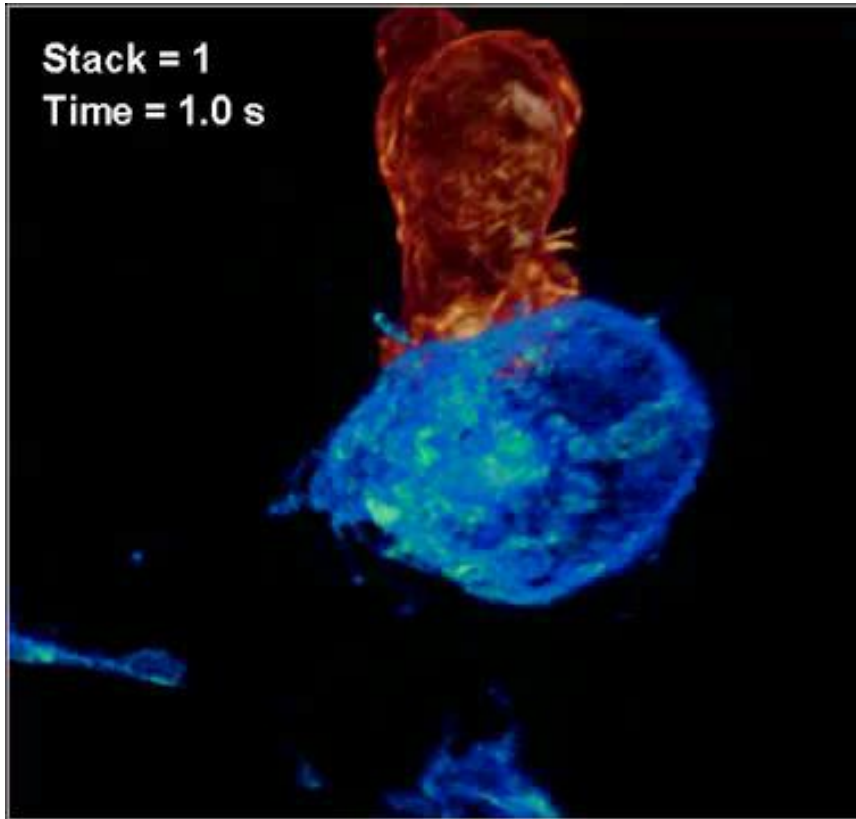
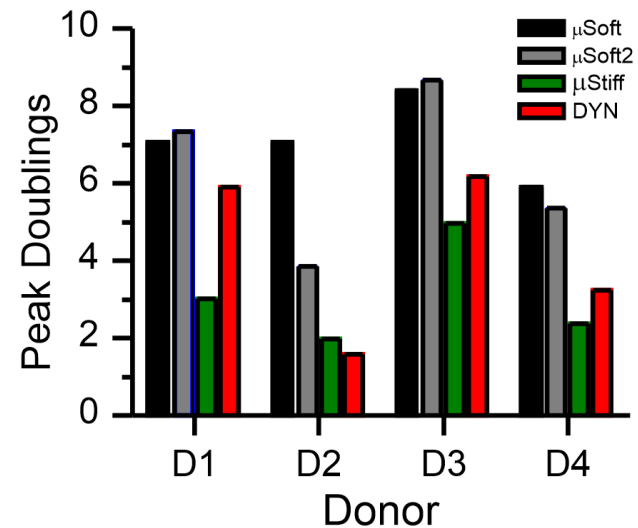
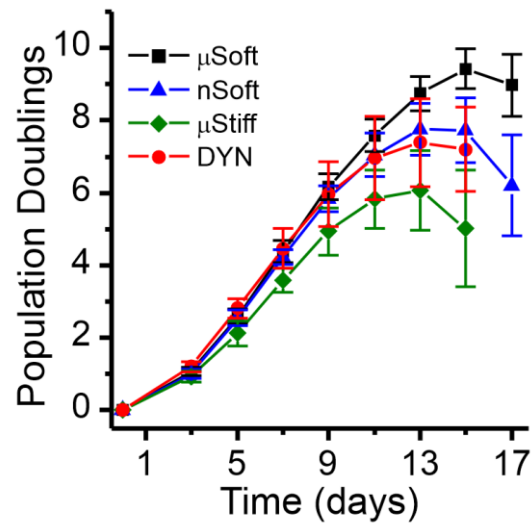
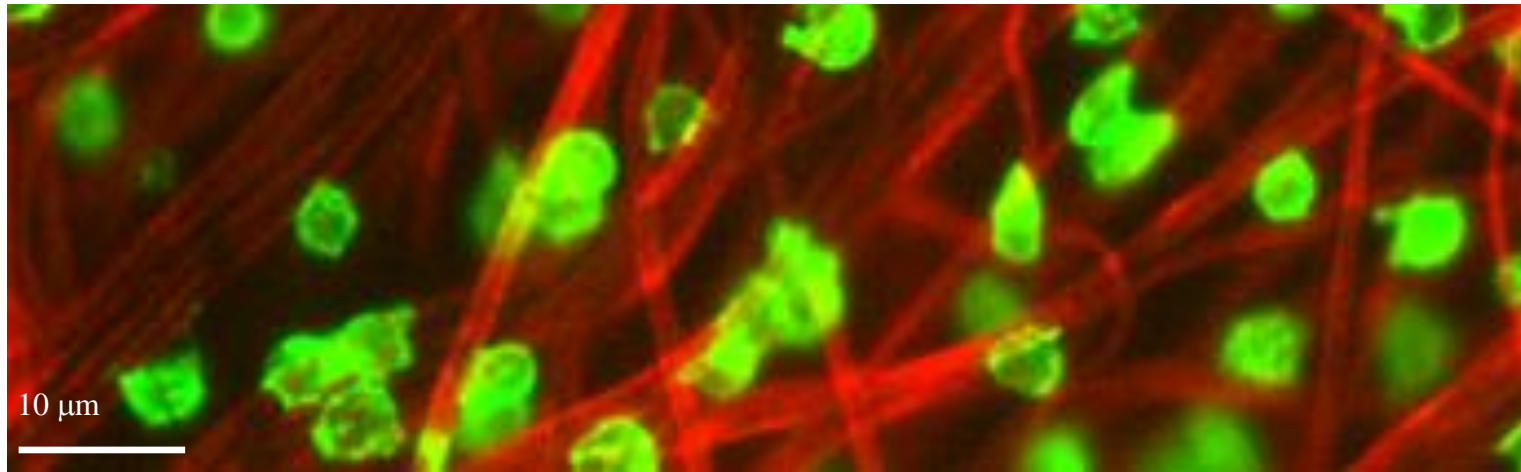


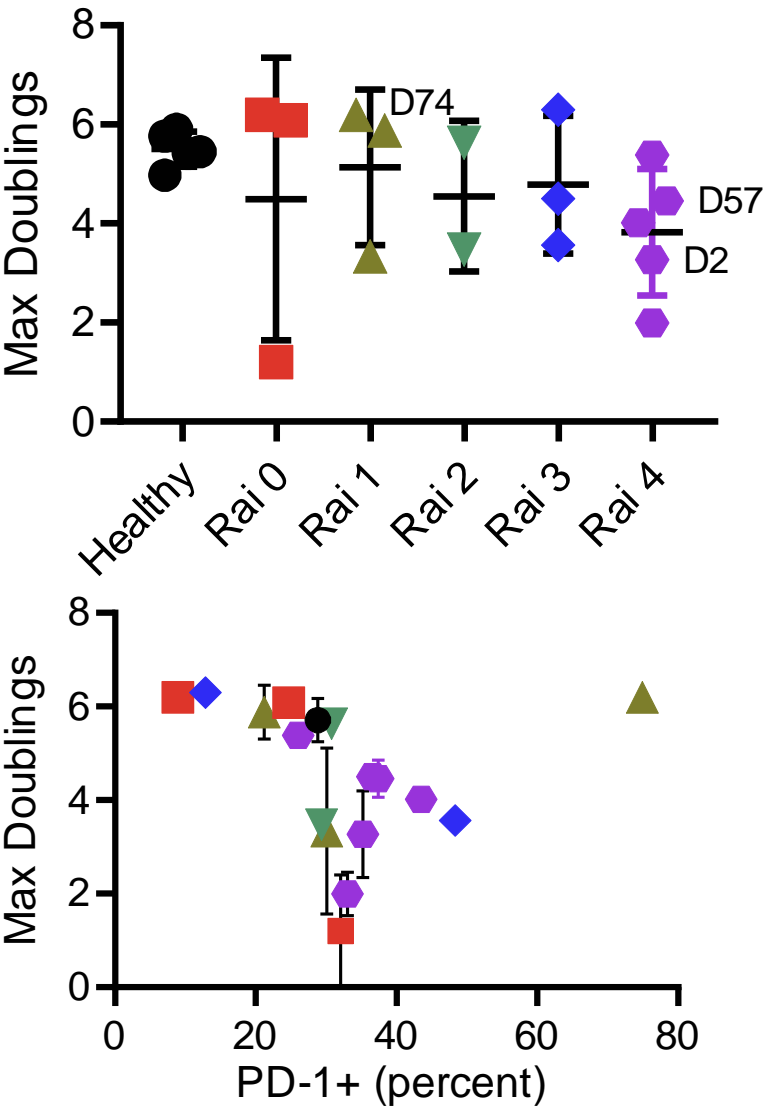
# Improving T cell expansion through mechanobiology



# Improving T cell expansion through mechanobiology

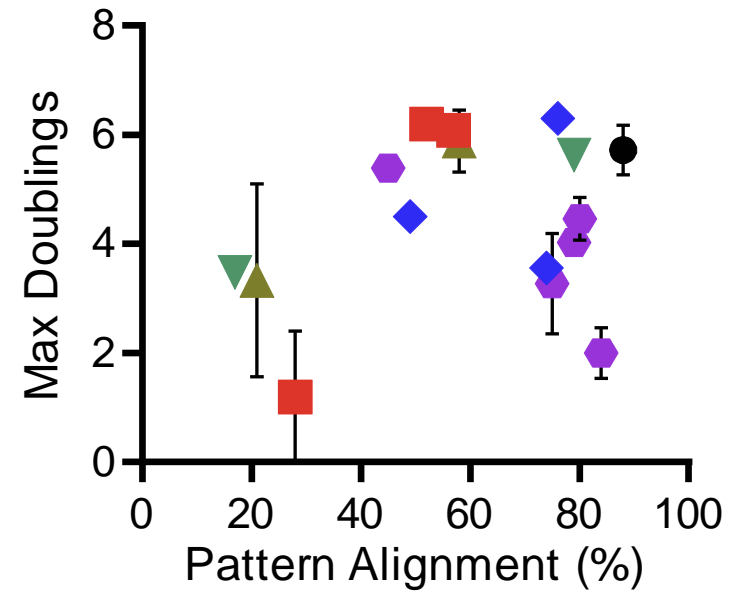
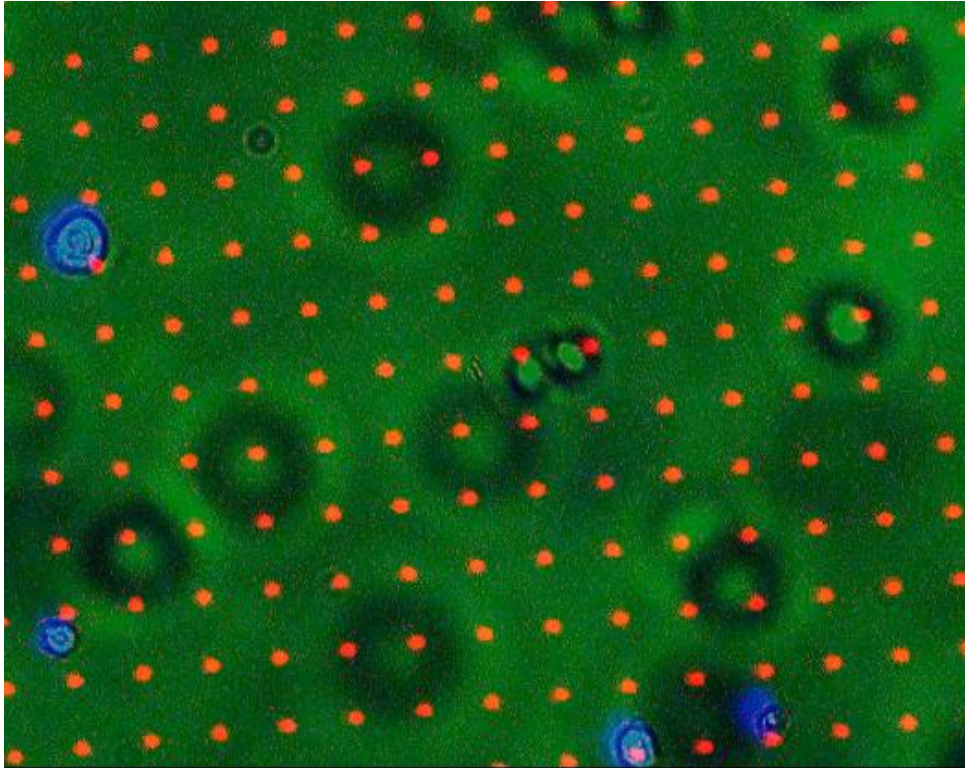


# Understanding variability in disease: biomarkers



Patient No	Age	IgVH	PD-1 (%)	Rai	Sex	IL-2	Max Doublings
D2	59	MUT	35.20	3	M	10.75	3.28
D44	81		74.88	1	M		6.20
D46	71		33.04	4	F	30.79	2.00
D47	71	MUT	56.59	4	F	11.26	
D48	56	UNMUT	45.59	1	M	14.62	
D54	63	UNMUT	26.03	4	F	110.09	5.39
D55	88	UNMUT	48.36	3	M	42.61	3.56
D57	84	MUT	37.38	4	M	14.79	4.46
D58	52	UNMUT	36.56	3	F	46.33	4.50
D59	45	MUT	8.29	3	F		
D62	42	UNMUT	43.50	4	M	36.74	4.02
D65	69	UNMUT	39.67	2	M	3.79	
D66	52	MUT	24.70	0	F	62.33	6.09
D68	83	MUT	12.85	3	M	90.53	6.30
D69	48		30.87	2	M		5.63
D74	59		21.19	1	F	66.98	5.88
D75	76	UNMUT	32.09	0	F	27.11	1.20
D76	57	MUT	29.35	2	F	14.48	3.48
D77	47	MUT	8.89	0	M	110.81	6.20
D78	69	MUT	30.07	1	F	25.91	3.34
H3							5.45
H4	34		19.00		F	283.4	4.98
H6	49		20.18		F	119.69	5.76
H8							5.42
H9	49		50.70		M	65.63	5.91

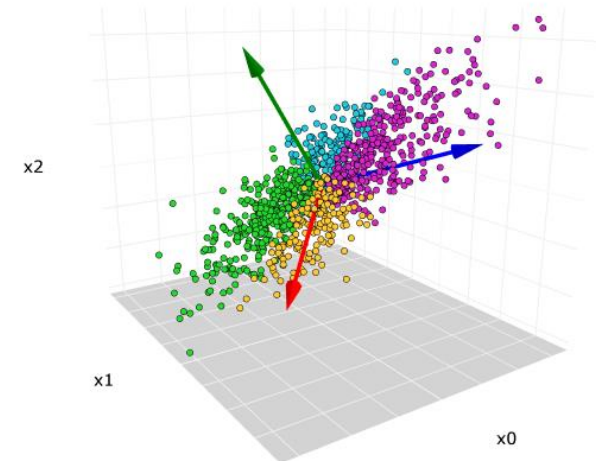
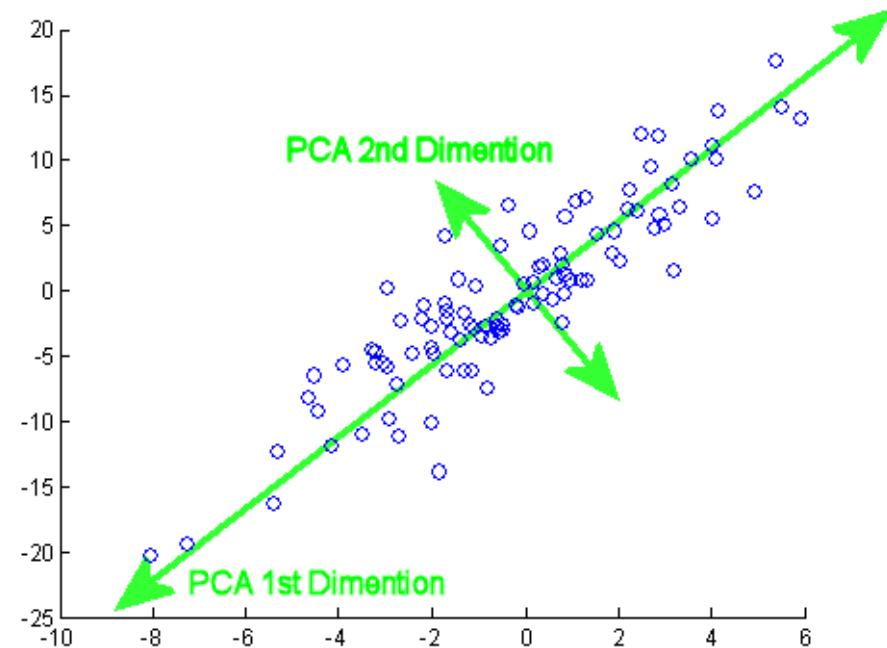
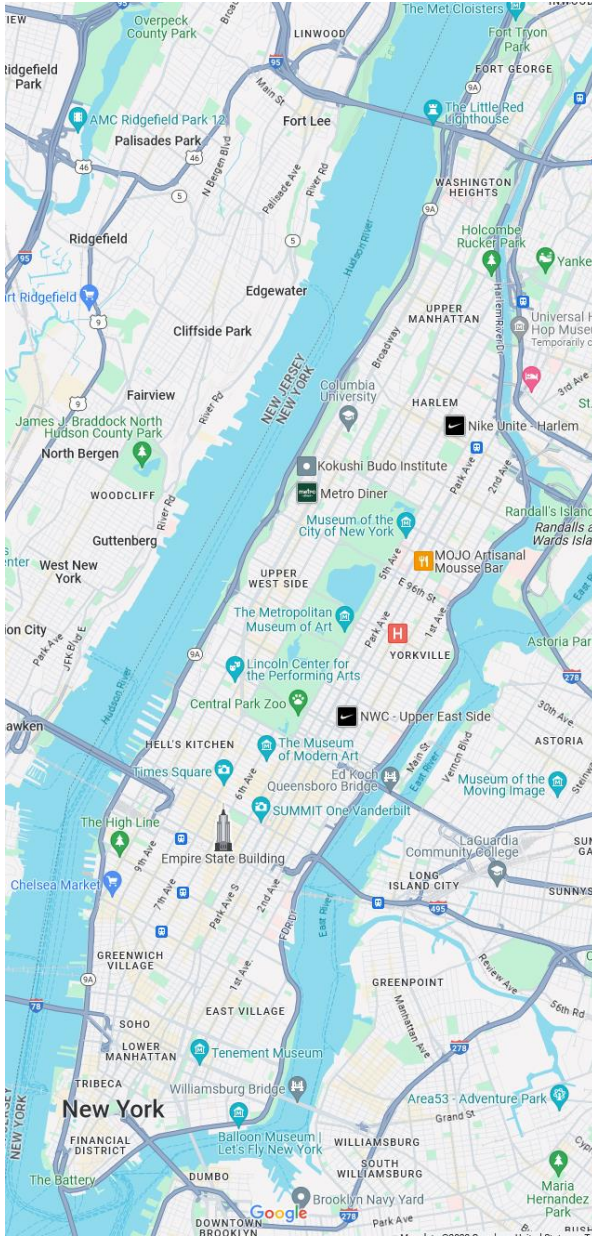
## Understanding variability in disease: measures of cell function



Patient No	Age	IgVH	PD-1 (%)	Rai	Sex	Pattern Alignment (%)	IL-2	Max Doublings
D2	59	MUT	35.20	4	M	75	10.75	3.28
D44	81	UNMUT*	74.88	1	M	75*	11.26*	6.20
D46	71	MUT*	33.04	4	F	84	30.79	2.00
D47	71	MUT	56.59	4	F	60	11.26	
D48	56	UNMUT	45.59	1	M	71	14.62	
D54	63	UNMUT	26.03	4	F	45	110.09	5.39
D55	88	UNMUT	48.36	3	M	74	42.61	3.56
D57	84	MUT	37.38	4	M	80	14.79	4.46
D58	52	UNMUT	36.56	3	F	49	46.33	4.50
D59	45	MUT	8.29	3	F	13	110.81*	
D62	42	UNMUT	43.50	4	M	79	36.74	4.02
D65	69	UNMUT	39.67	2	M	32	3.79	
D66	52	MUT	24.70	0	F	57	62.33	6.09
D68	83	MUT	12.85	3	M	76	90.53	6.30
D69	48	MUT*	30.87	2	M	79	46.33*	5.63
D74	59	MUT*	21.19	1	F	58	66.98	5.88
D75	76	UNMUT	32.09	0	F	28	27.11	1.20
D76	57	MUT	29.35	2	F	17	14.48	3.48
D77	47	MUT	8.89	0	M	52	110.81	6.20
D78	69	MUT	30.07	1	F	21	25.91	3.34
H3								5.45
H4	34		19.00		F	94	283.4	4.98
H6	49		20.18		F	85	119.69	5.76
H8								5.42
H9	49		50.70		M	86	65.63	5.91

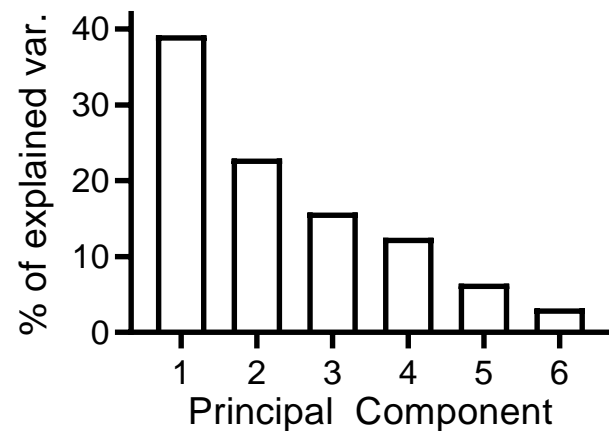
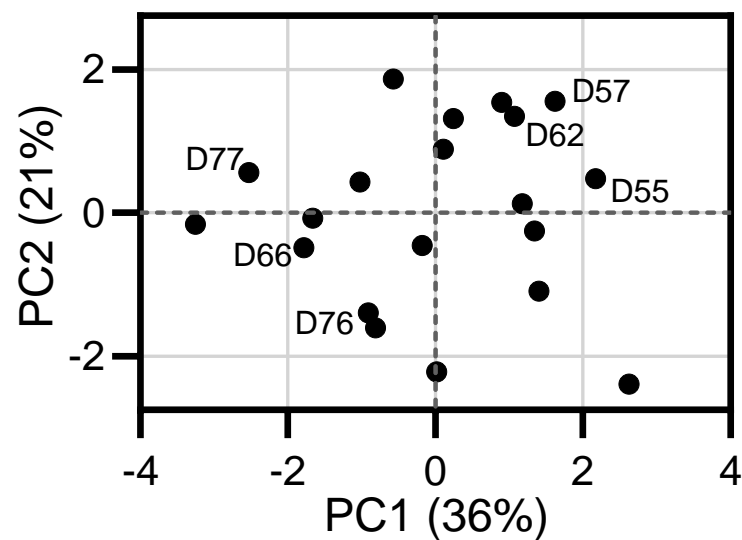


# Dealing with multiple variables. Principal Component Analysis



# PCA of expansion data

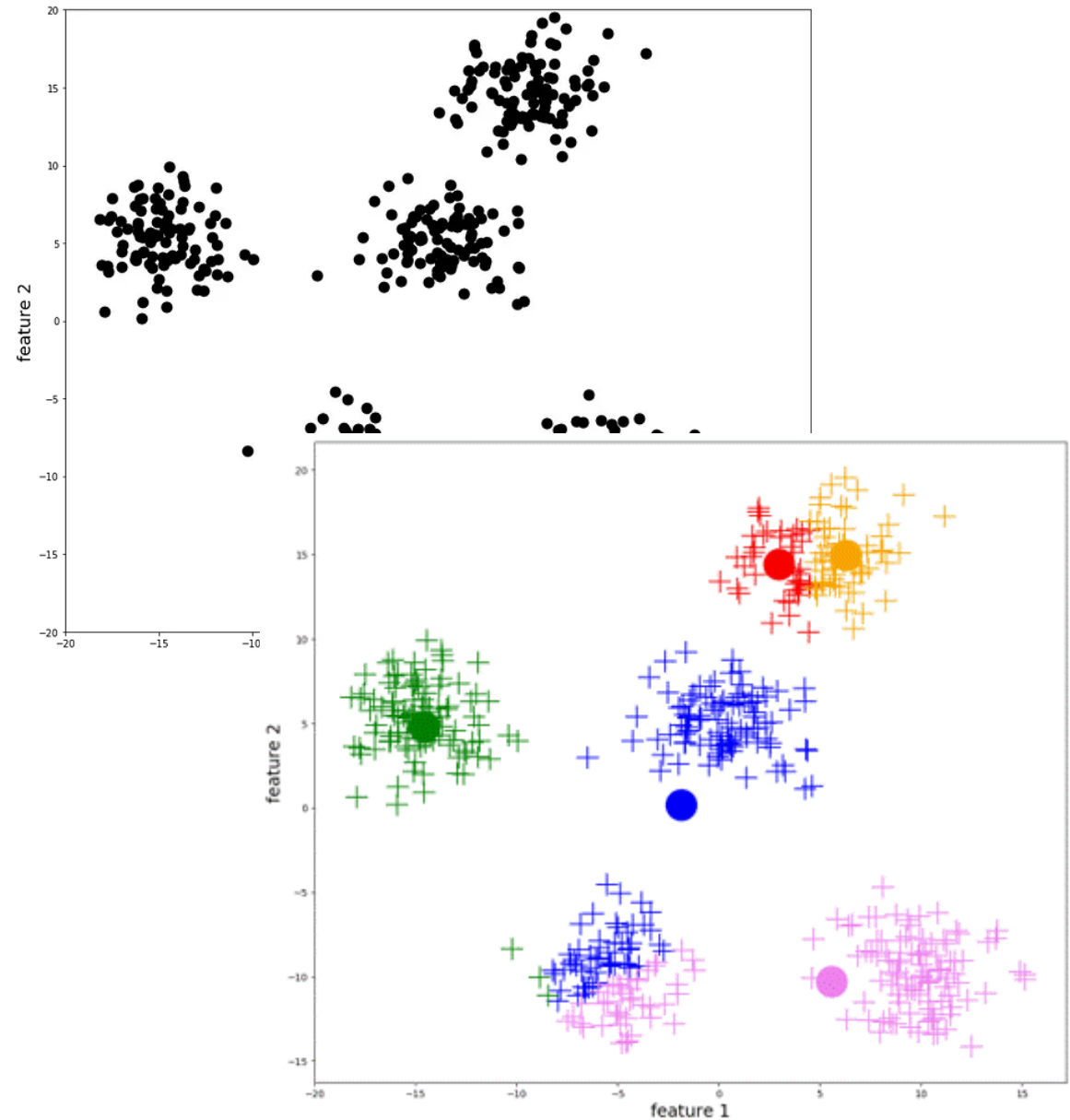
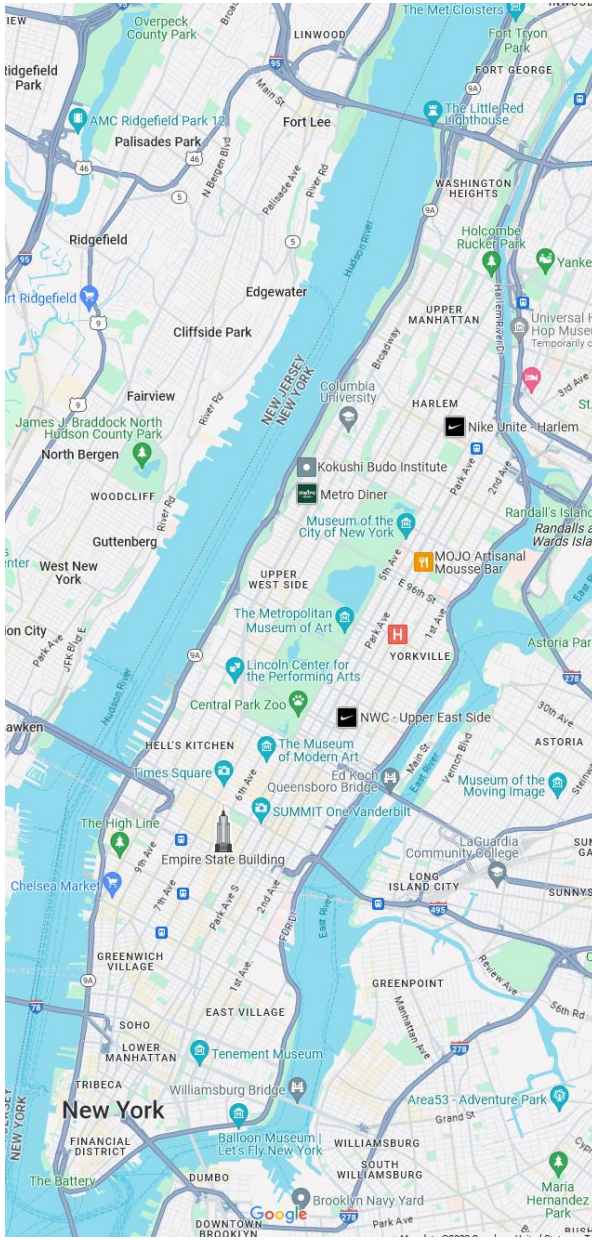
Patient No	Age	IgVH	PD-1 (%)	Rai	Sex	Pattern Alignment (%)	IL-2	Max Doublings
D2	59	MUT	35.20	4	M	75	10.75	3.28
D44	81	UNMUT*	74.88	1	M	75*	11.26*	6.20
D46	71	MUT*	33.04	4	F	84	30.79	2.00
D47	71	MUT	56.59	4	F	60	11.26	
D48	56	UNMUT	45.59	1	M	71	14.62	
D54	63	UNMUT	26.03	4	F	45	110.09	5.39
D55	88	UNMUT	48.36	3	M	74	42.61	3.56
D57	84	MUT	37.38	4	M	80	14.79	4.46
D58	52	UNMUT	36.56	3	F	49	46.33	4.50
D59	45	MUT	8.29	3	F	13	110.81*	
D62	42	UNMUT	43.50	4	M	79	36.74	4.02
D65	69	UNMUT	39.67	2	M	32	3.79	
D66	52	MUT	24.70	0	F	57	62.33	6.09
D68	83	MUT	12.85	3	M	76	90.53	6.30
D69	48	MUT*	30.87	2	M	79	46.33*	5.63
D74	59	MUT*	21.19	1	F	58	66.98	5.88
D75	76	UNMUT	32.09	0	F	28	27.11	1.20
D76	57	MUT	29.35	2	F	17	14.48	3.48
D77	47	MUT	8.89	0	M	52	110.81	6.20
D78	69	MUT	30.07	1	F	21	25.91	3.34
H3								5.45
H4	34		19.00		F	94	283.4	4.98
H6	49		20.18		F	85	119.69	5.76
H8								5.42
H9	49		50.70		M	86	65.63	5.91





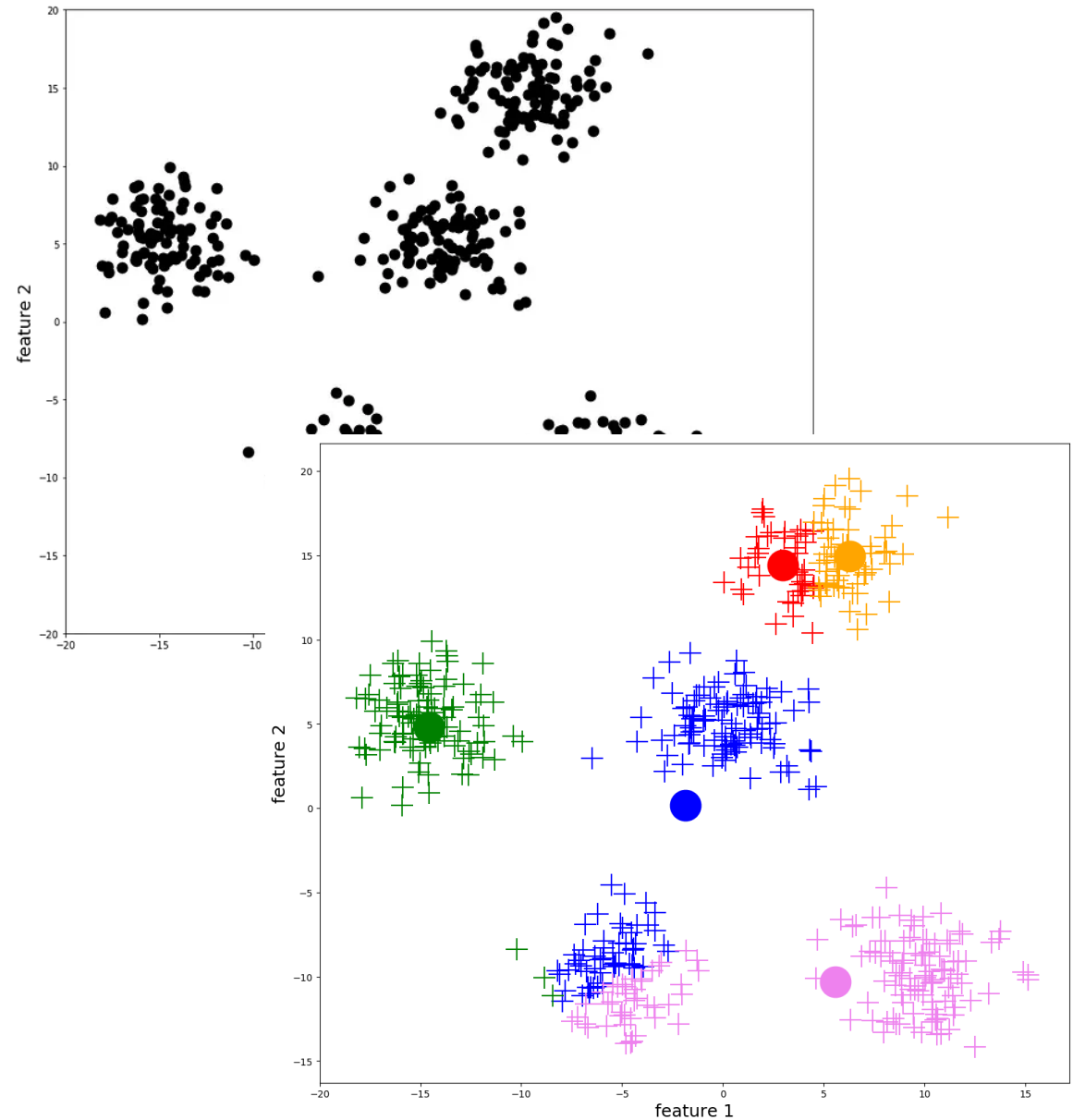
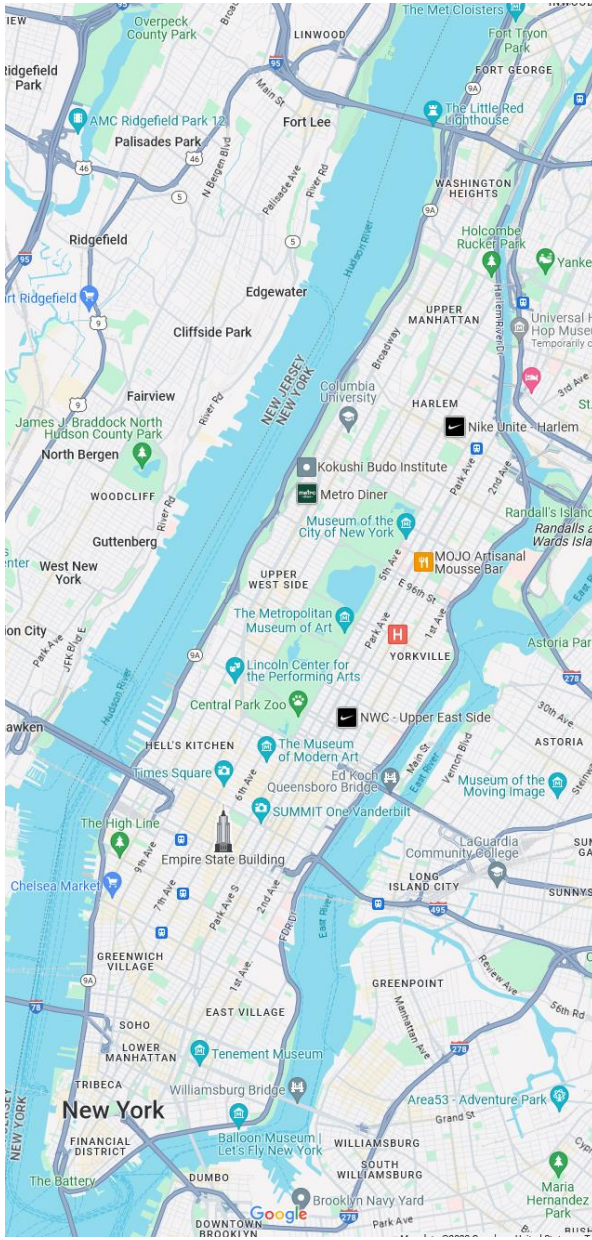


# Finding patterns: k-means clustering

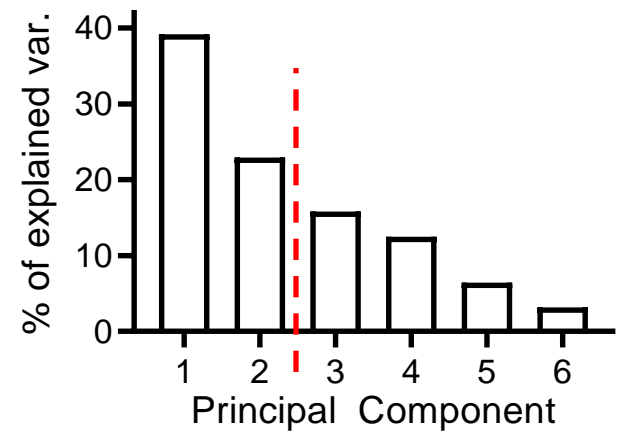
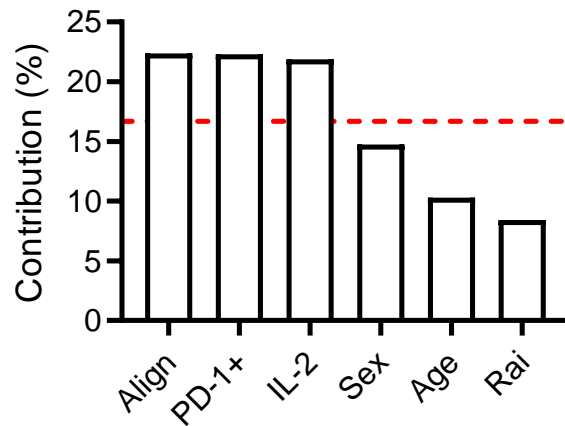
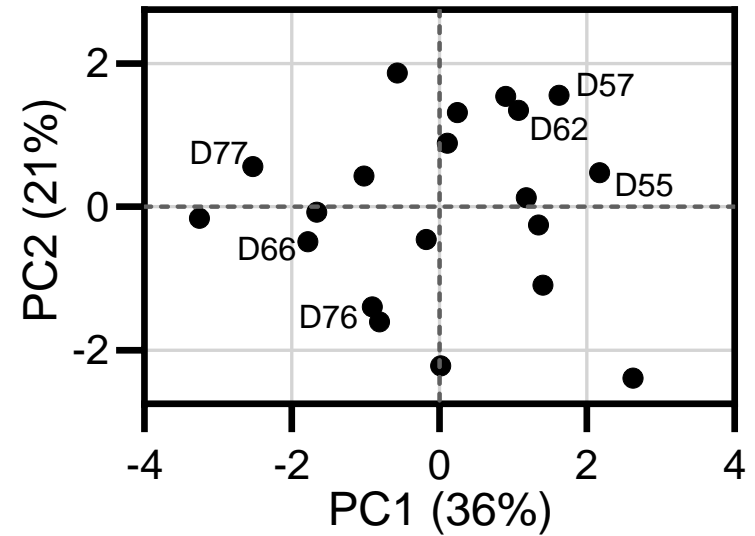
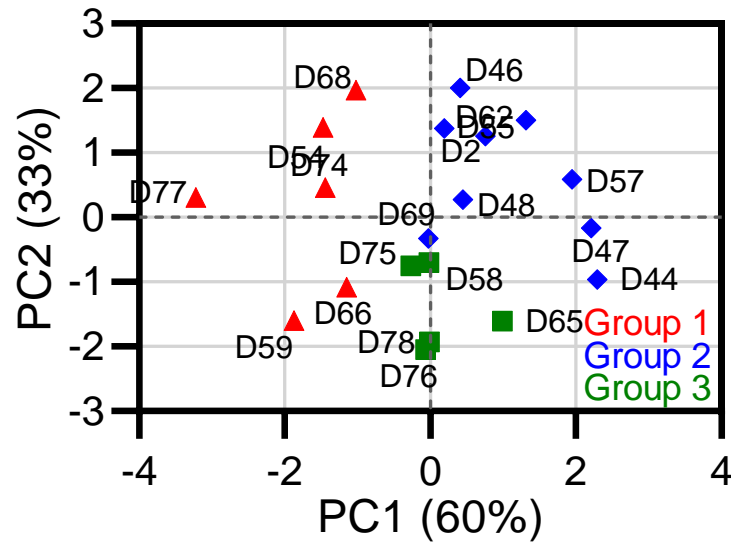




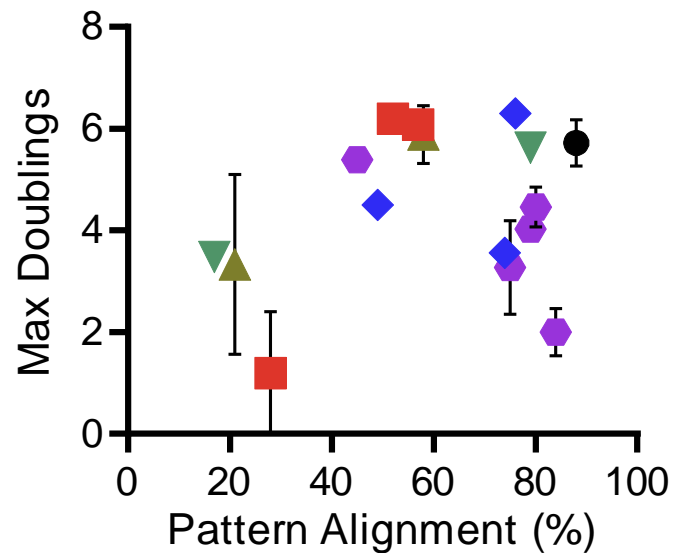
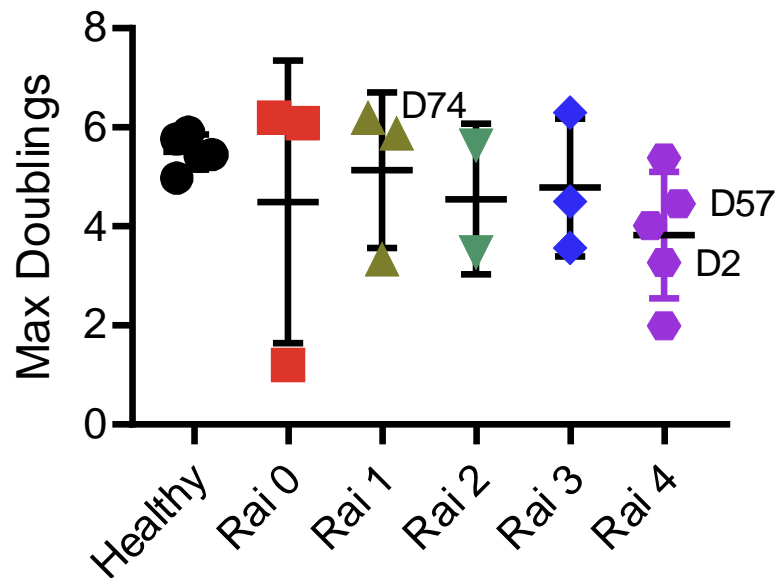
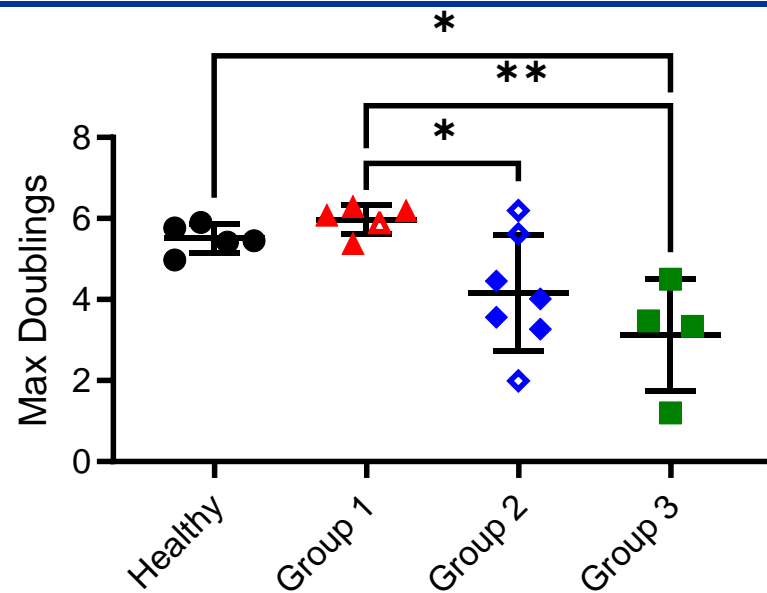
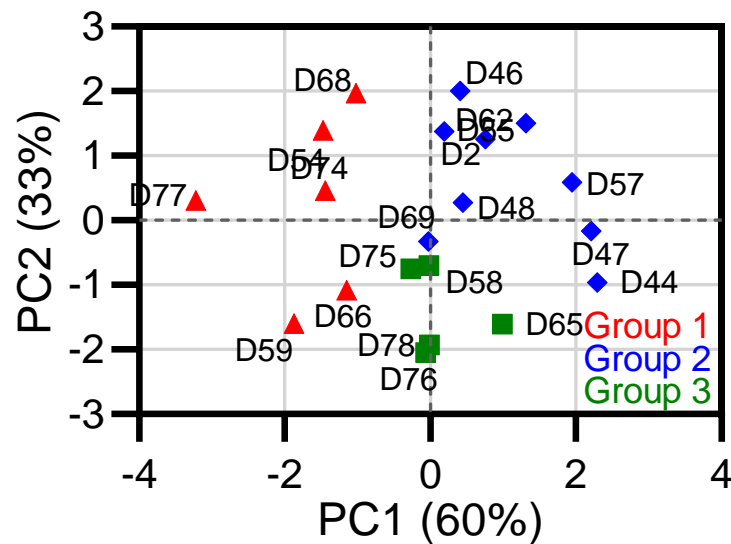
# Finding patterns: k-means clustering



## PCA of expansion data

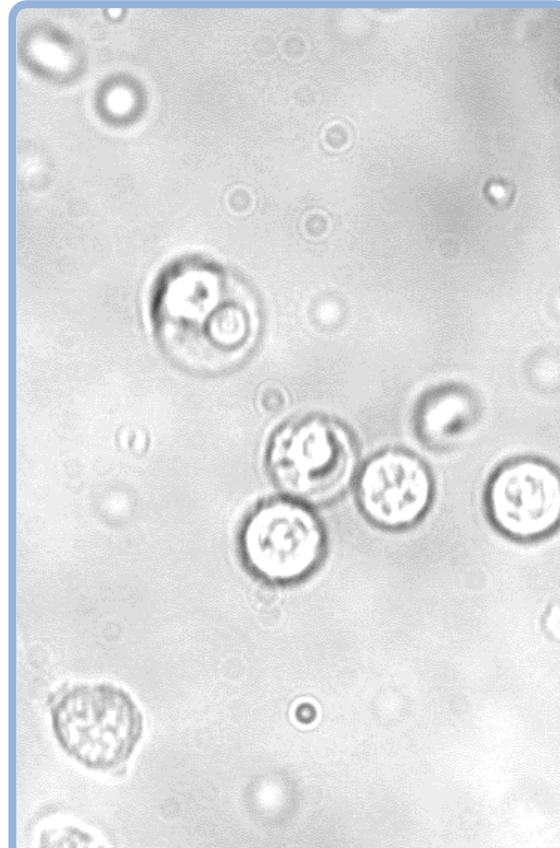


## PCA of expansion data



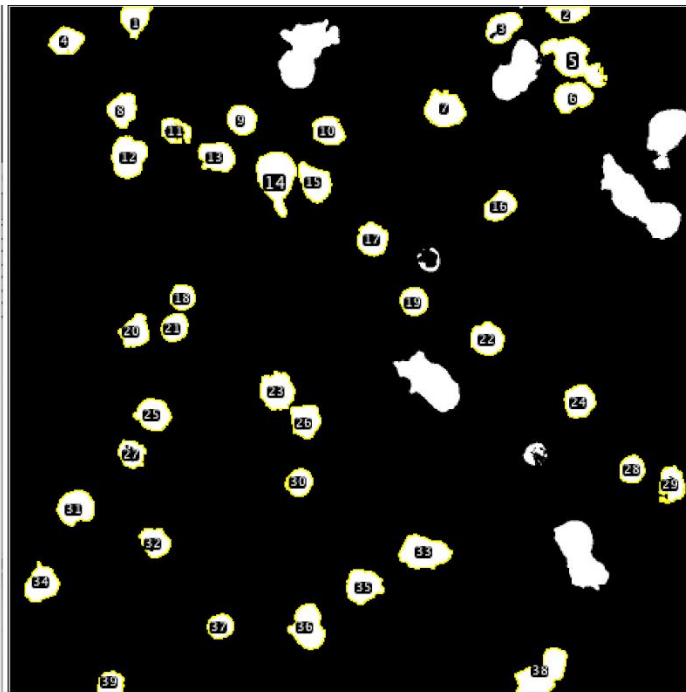
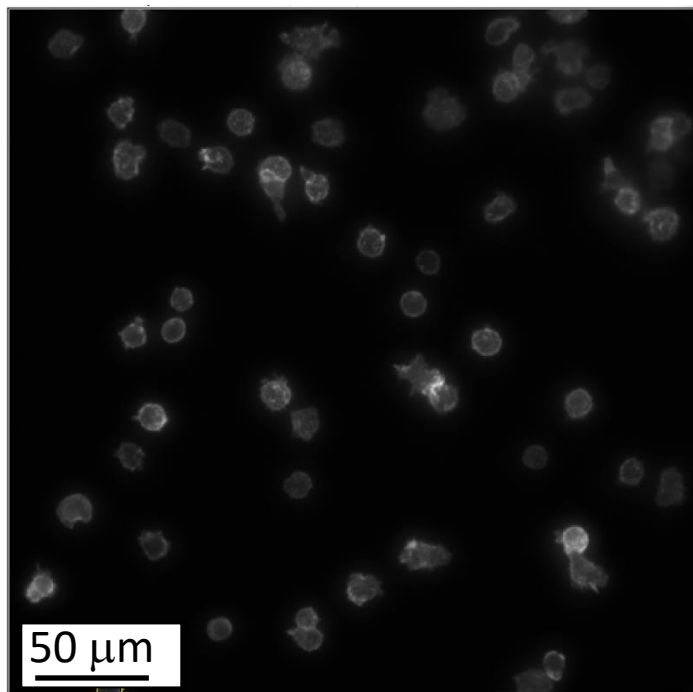


Healthy

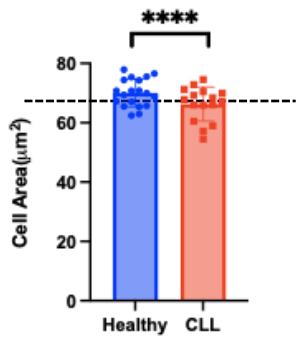
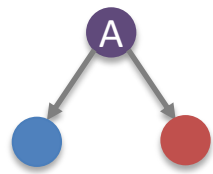


CLL



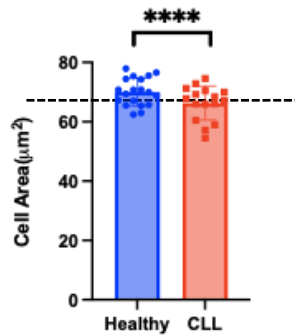
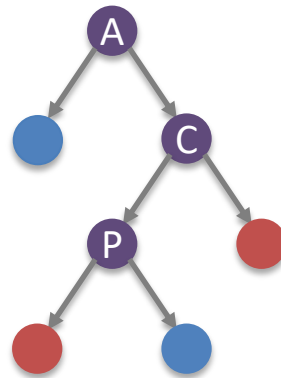


Morphology features
Area
Roundness
Major
Minor
Perimeter
Solidity
Circularity
Height
Width
Feret's Diameter
Aspect Ratio



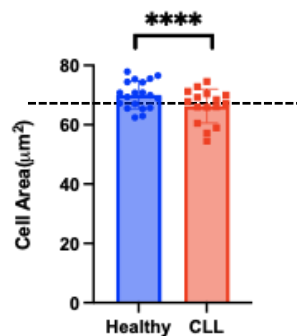
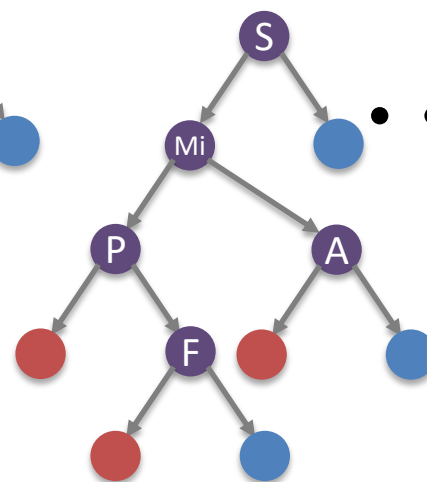
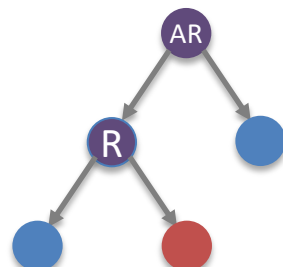
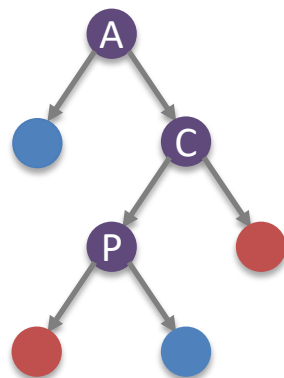
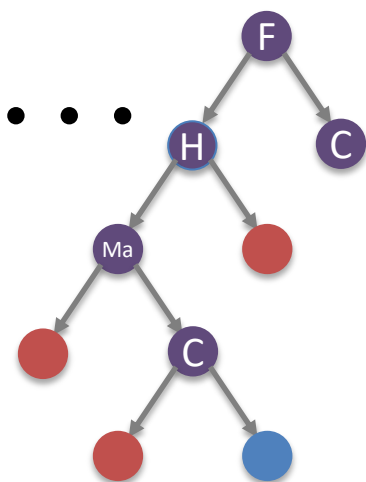
	Single-Feature (Area) Decision Tree
Accuracy	0.50
AUC	0.51
Sensitivity	0.43
Specificity	0.60

Morphology features
Area
Roundness
Major
Minor
Perimeter
Solidity
Circularity
Height
Width
Feret's Diameter
Aspect Ratio



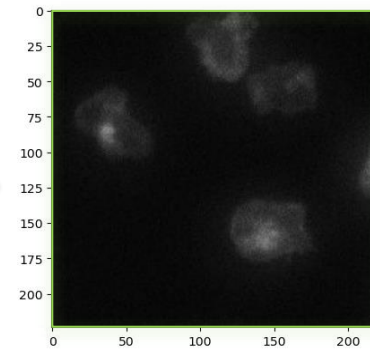
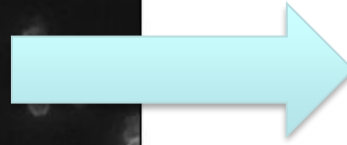
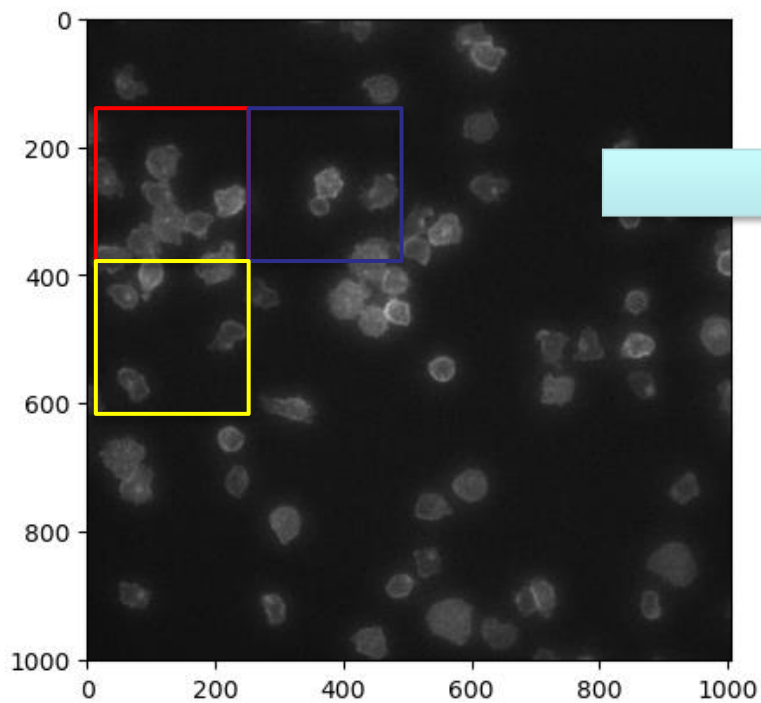
	Single-Feature (Area) Decision Tree	Multi-Feature Decision Tree
Accuracy	0.50	0.70
AUC	0.51	0.69
Sensitivity	0.43	0.79
Specificity	0.60	0.58

Morphology features
Area
Roundness
Major
Minor
Perimeter
Solidity
Circularity
Height
Width
Feret's Diameter
Aspect Ratio

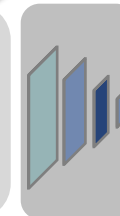


	Single-Feature (Area) Decision Tree	Multi-Feature Decision Tree	Random Forest
Accuracy	0.50	0.70	0.71
AUC	0.51	0.69	0.83
Sensitivity	0.43	0.79	0.79
Specificity	0.60	0.58	0.60

Morphology features
Area
Roundness
Major
Minor
Perimeter
Solidity
Circularity
Height
Width
Feret's Diameter
Aspect Ratio



Feature Extractor



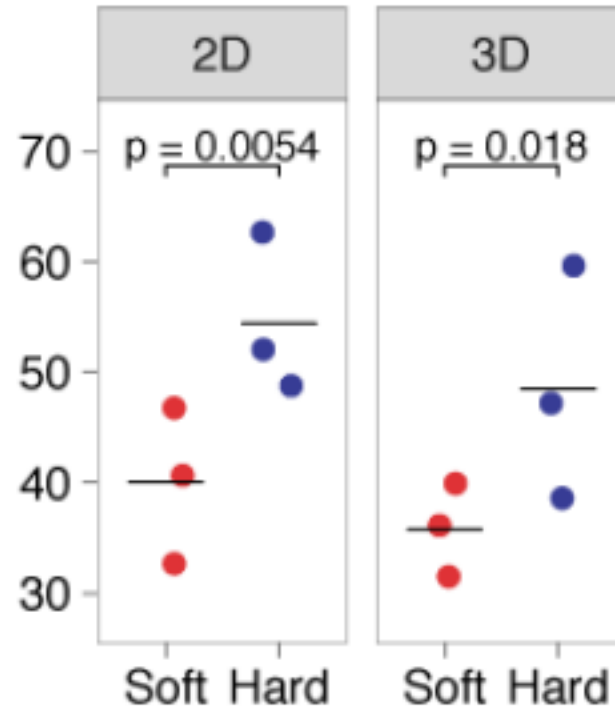
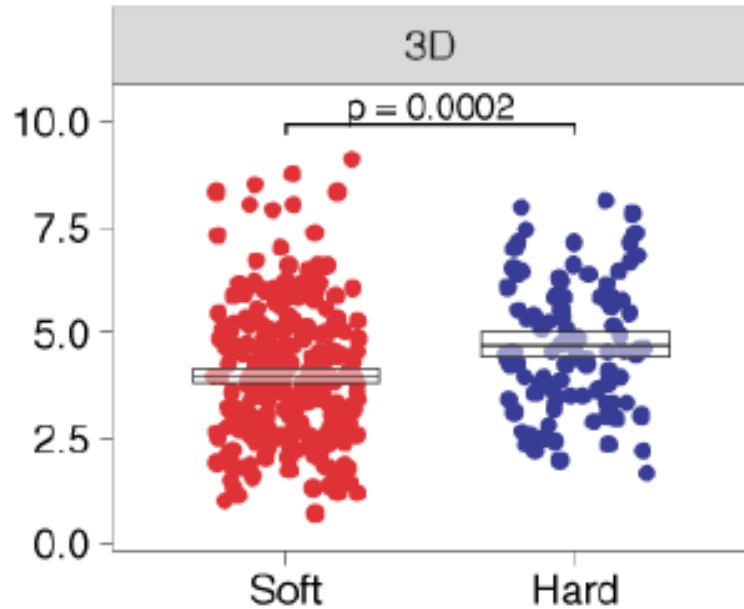
Classifier



**0: Healthy**

**1: CLL**

	Machine Learning			Deep Learning			
	Single-Feature (Area) Decision Tree	Multi-Feature Decision Tree	Random Forest	Pretrained Swin Transformer Unfrozen weights	Pretrained Swin Transformer Frozen weights	Un-pretrained Swin Transformer	Pretrained Resnet 50
Accuracy	0.538	0.643	0.693	0.859	0.720	0.754	0.817
AUC	0.540	0.641	0.768	0.887	0.887	0.783	0.845
Sensitivity	0.559	0.681	0.724	0.810	0.642	0.721	0.801
Specificity	0.516	0.601	0.659	0.929	0.833	0.801	0.840

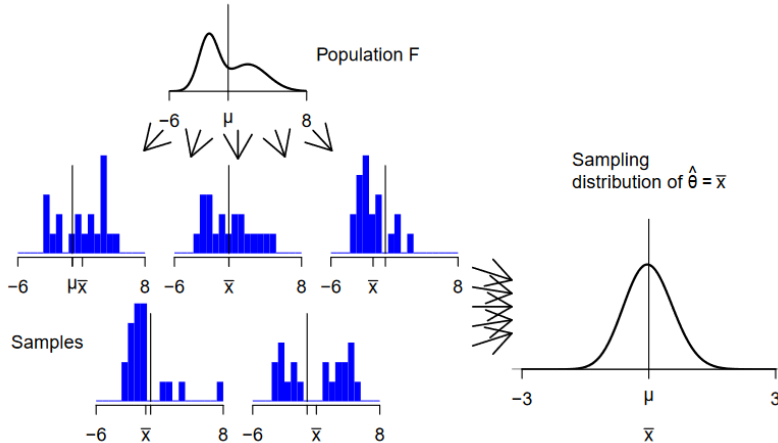


Bootstrapping confidence intervals on the mean

- “What Teachers Should Know About the Bootstrap:”  
<https://arxiv.org/pdf/1411.5279.pdf>

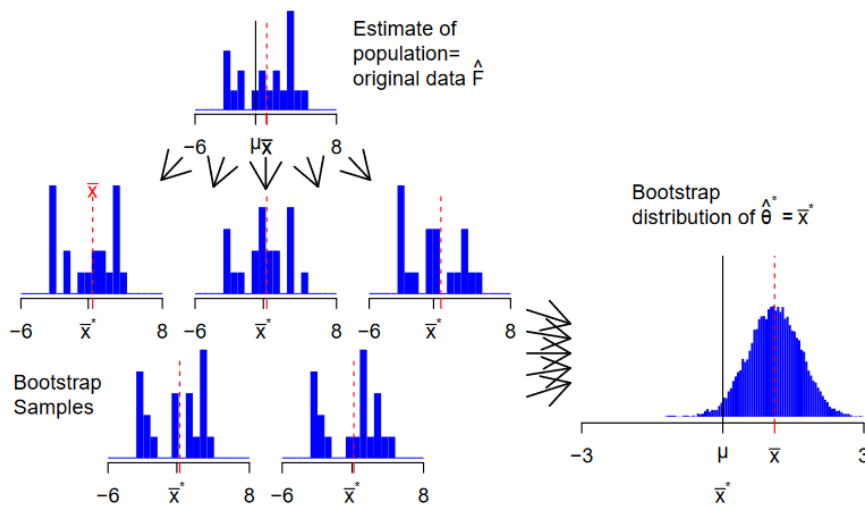


# Bootstrapping



## Ideal

- Draw samples from the population
- Compute statistic from sample, making a sampling distribution



## Bootstrapping

- Draw samples from *an estimate of the population*
- Compute statistic of interest from each sample
- Distribution of statistics is the bootstrap distribution

# Permutation Test

Basic	6.95	10.013	10.62	10.15	8.583
	7.62	8.233	10.35	11.016	8.516
Extended	3.383	7.8	9.416	4.66	5.36
	7.63	4.95	8.013	7.8	9.58

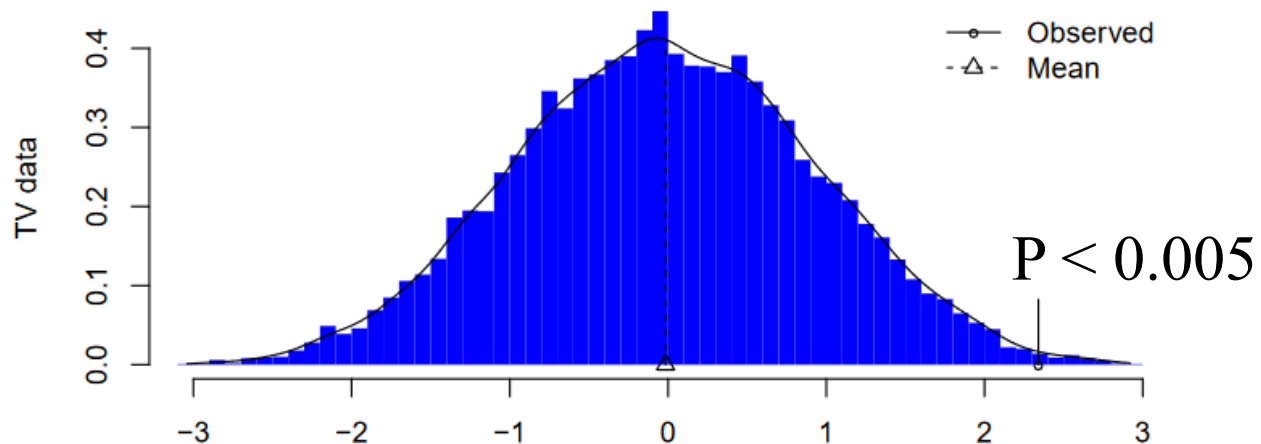
Table 1: *Minutes of commercials per half-hour of TV.*

t-Test: Two-Sample Assuming Unequal Variances		
	<i>Basic</i>	<i>Extended</i>
Mean	9.2051	6.8592
Variance	1.949167	4.494296
Observations	10	10
Hypothesized Mean Difference	0	
df	16	
t Stat	2.922468	
P(T<=t) one-tail	0.004983	
t Critical one-tail	1.745884	
P(T<=t) two-tail	0.009965	
t Critical two-tail	2.119905	

# Permutation Test

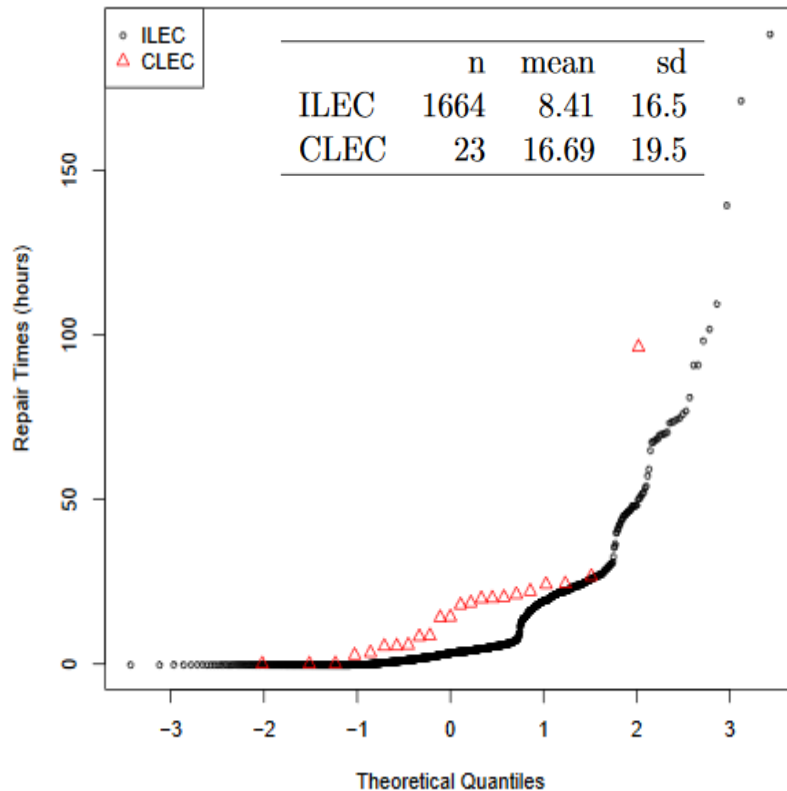
Basic	6.95	10.013	10.62	10.15	8.583
	7.62	8.233	10.35	11.016	8.516
Extended	3.383	7.8	9.416	4.66	5.36
	7.63	4.95	8.013	7.8	9.58

- Pool all 20 observations, pick 10 as Basic, 10 as Extended, without replacement
- Generate a permutation distribution around a statistic
  - Difference in means
  - Compute this statistic for a given pick.
  - Repeat pick many times to generate distribution
  - Compare observed to simulated distribution



# Case for resampling

See page 37 of “What teachers...”



- Verizon is to provide repairs for CLEC (competitors) as quickly as for ILEC (theirs).
- If  $CLEC > ILEC$ , this would bring penalties.
- Criteria for penalties was 1% significance level.

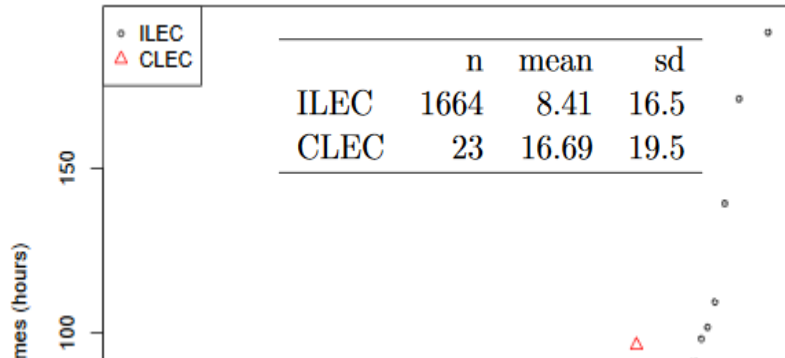
---

	$t$	$P$ -value
Permutation test		0.0171
Pooled $t$ test	2.61	0.0045
PUC $t$ test	2.63	0.0044

---

# Case for resampling

See page 37 of “What teachers...”



- Verizon is to provide repairs for CLEC (competitors) as quickly as for ILEC (theirs).
- If  $ILEC > CLEC$ , this would

So, given the discrepancy between the permutation test result and the various  $t$  tests, which one is right? Absolutely, definitely, the permutation test. Sir Ronald Fisher originally argued for  $t$  tests by describing them as a computationally-feasible approximation to permutation tests (known to be the right answer), given the computers of the time. We should not be bound by that limitation.

	$t$	$P$ -value
Permutation test		0.0171
Pooled $t$ test	2.61	0.0045
PUC $t$ test	2.63	0.0044