

The Pulmonologist Perspective:
Probability of Malignancy is
determined by integration of
clinical and imaging features.

Charles A. Powell, M.D.

Chief, Pulmonary, Critical Care, and Sleep Medicine
Mount Sinai School of Medicine

Prior Probability of Malignancy: Enter a number from 1 to 100 %

Clinical Characteristics

Age:

Smoking (Pk-yrs):

Hemoptysis:

Hx Prev Malig:

Radiographic Characteristics

Size (cm):

Location:

Edge:

Growth Rate:

Cavity Wall Thickness:

Calcification:

Additional Characteristics

Contrast Enhancement:

PET:

The Probability of Malignancy is:

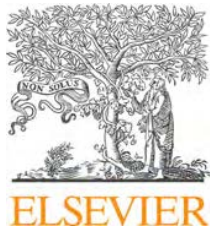
Likelihood Ratios

20-29	0.05
30-39	0.24
40-49	0.94
50-59	1.90
60-69	2.64
Nonsmoker	0.15
< 30 pk-yrs	0.74
30-39 pk-yrs	2
>40 pk-yrs	3.7
Hemoptysis, absent	1
Hemoptysis, present	5.08
No prev malig	1
Prev Malig	4.95
0-1 cm	0.52
1.1 - 2.0	0.74
2.1 - 3.0	3.67
> 3.0 cm	5.23
upper/middle	1.22
Lower	0.66
Smooth	0.3
Lobulated	0.74
Spiculated	5.54
Growth, not known	1
Benign growth rate	0.01
Malignant growth rate	3.4
Not cavitated	1
< 4 mm	0.07
5 - 15 mm	0.72
> 16	38
Not calcified	2.2
Benign calcification	0.01
Enhancement < 15 HU	0.04
Enhancement > 15 HU	2.32
SUR < 2.5	0.06
SUR > 2.5	7.1

What Tool Characteristics Will be Most Helpful

Automated Tool that functions as “second reader”

Features designed to maximize sensitivity, such that the test is analogous the D-Dimer assay for diagnosis of Pulmonary Embolism (a negative test essentially rules out disease)



COLUMBIA UNIVERSITY
MEDICAL CENTER



Computer-Aided Diagnosis (CADx) for Lung Cancer

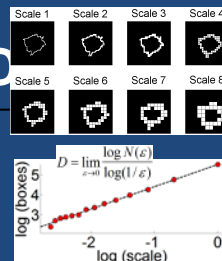
Computer-Aided Diagnosis (CADx)

- the next step after detection
- aims at the **characterization** of abnormalities in medical images
- uses image analysis and machine learning for evidence integration to help **determine whether or not a lesion is cancerous**

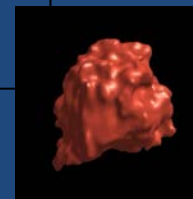
194 image-based features

56 2D + 56 2½D + 82 3D

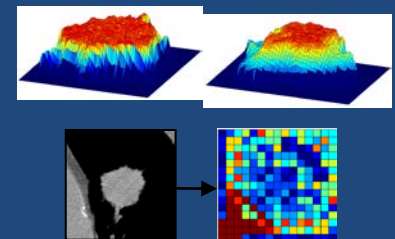
fractal analysis



shape analysis



texture



22 clinical features

- basic patient demographics (age, gender)
- clinical risk factors (prior chest surgery, emphysema, lymph node status, satellite nodules)



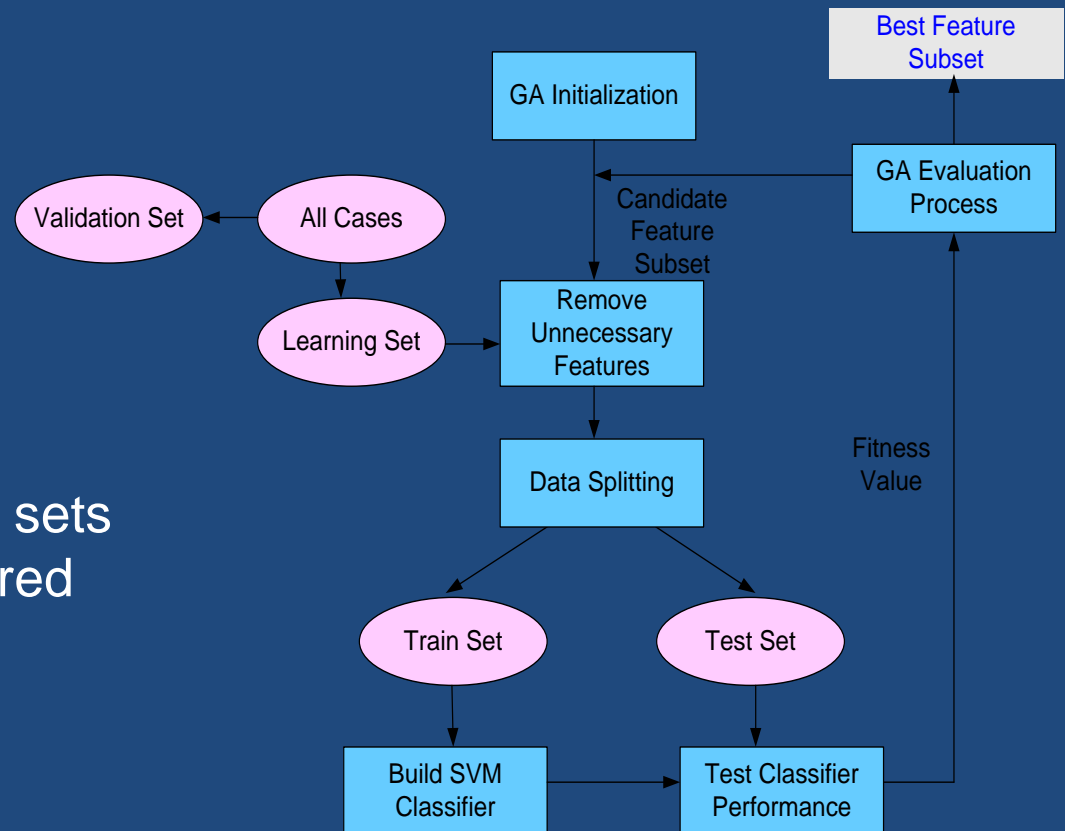
Feature Selection – Genetic Algorithm

Identify the truly relevant features for diagnosis:

- Improve accuracy of classifier
- More computationally efficient

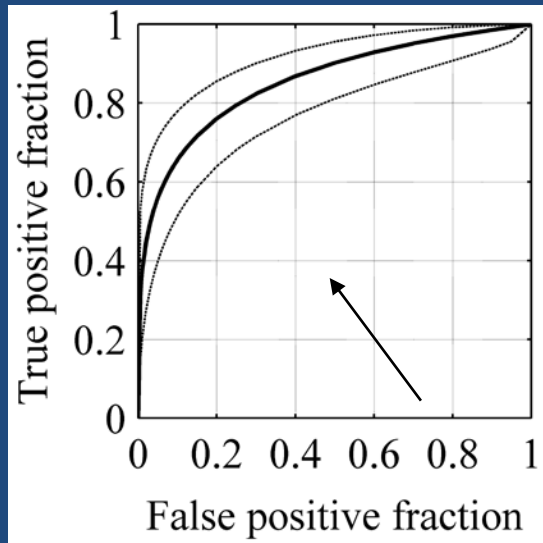
Genetic Algorithm

- Optimization technique that iteratively evaluates feature sets
- Optimal feature set discovered by evolutionary process



Boroczky L "Feature subset selection for improving the performance of false positive reduction in lung nodule CAD", *IEEE Trans Inf Technol Biomed.* 2006; 10(3):504–11.

Validation (leave one out) n=128



**Leave-one-out validation using
two-step Genetic Algorithm-
ensemble SVM ensemble
classifier**

sensitivity: 90%

specificity: 71%

PPV: 76%

NPV: 88%

Az: 0.85 (0.77-0.90 C.I.)

Conclusion

- A CADx system based on an SVM ensemble has been proposed
- Combined 2D and 3D analysis can give robustness to scan parameters
- Results suggest the advantage of ensemble classifiers over individual classifiers in achieving a reliable prediction of malignancy of pulmonary nodules
- Components can be generalized beyond lung cancer diagnosis

