Biomarkers for ovarian cancer

Mentee: Yuning Hu
Mentor: Antonio Olivas

If we don’t have any symptom, how would we know whether we’re having cancer?

We need a screening test!

Biomarkers

- Biological markers
- Continuous value
- Objective indications of medical state
- Can be measured accurately and reproducibly

- E.g. hemoglobin A1c (HbA1c) for diabetes

doi: 10.1097/COH.0b013e32833ed177

Binary test

	Disease	Non-disease
Test: Positive	True Positive	False Positive
Test: Negative	False Negative	True Negative

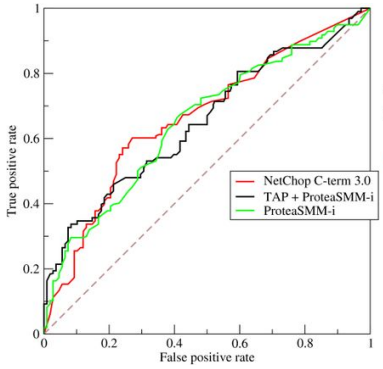
- Sensitivity = TP/(TP+FN)
- Specificity = TN/(TN + FP)
- Null hypothesis: not having the disease
- Alternative hypothesis: having the disease
- 1 - Sensitivity = Type 2 Error
- 1 - Specificity = Type 1 Error

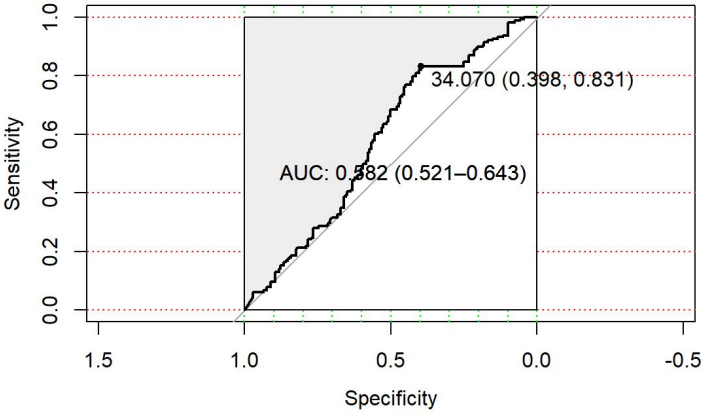
Test positive -> predict having the disease
Test negative -> predict not having the disease

How good a biomarker is?

ROC

- Receiver operating characteristic
- graphical plot that illustrates the performance of a biomarker at varying threshold values
- Fix a cutoff -> binary test
- True positive rate v.s. False positive rate
- Sensitivity v.s. 1 - Specificity





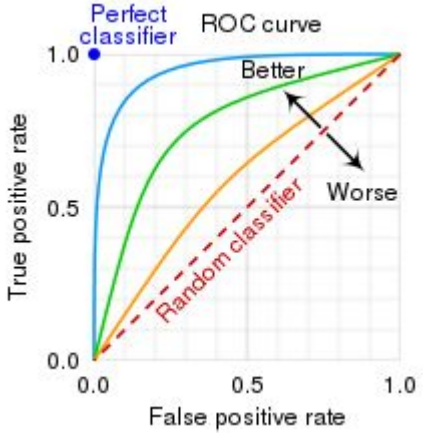
Wikipedia

If more than 1 biomarker, which one is better?

AUC

- area under the curve
- mathematically, it's integration
- expected value of performance of a biomarker
- higher AUC -> averagely, better performance

- Useless TPF = FPF
- => AUC = 0.5



Wikipedia

If we have a biomarker,
how to get the threshold
differentiating
disease & non-disease?

Cutoff

- ROC: cutoff
- Standard approach:
- penalize FNF (Type 2 Error), FPF (Type 1 Error) equally
- Minimizing (1 - Sensitivity) +(1 - Specificity)
- => Minimizing (Type 2 Error) +(Type 1 Error)

Would you rather
Have a disease but doctor says you don’t, or
Don’t have a disease but doctor says you have

Why modify the method for finding cutoff?

- Minimizing 1.5* (1 - Sensitivity) + (1 - Specificity)
- => Minimizing 1.5* (Type 2 Error) + (Type 1 Error)

Ovarian Cancer

Ovarian Cancer

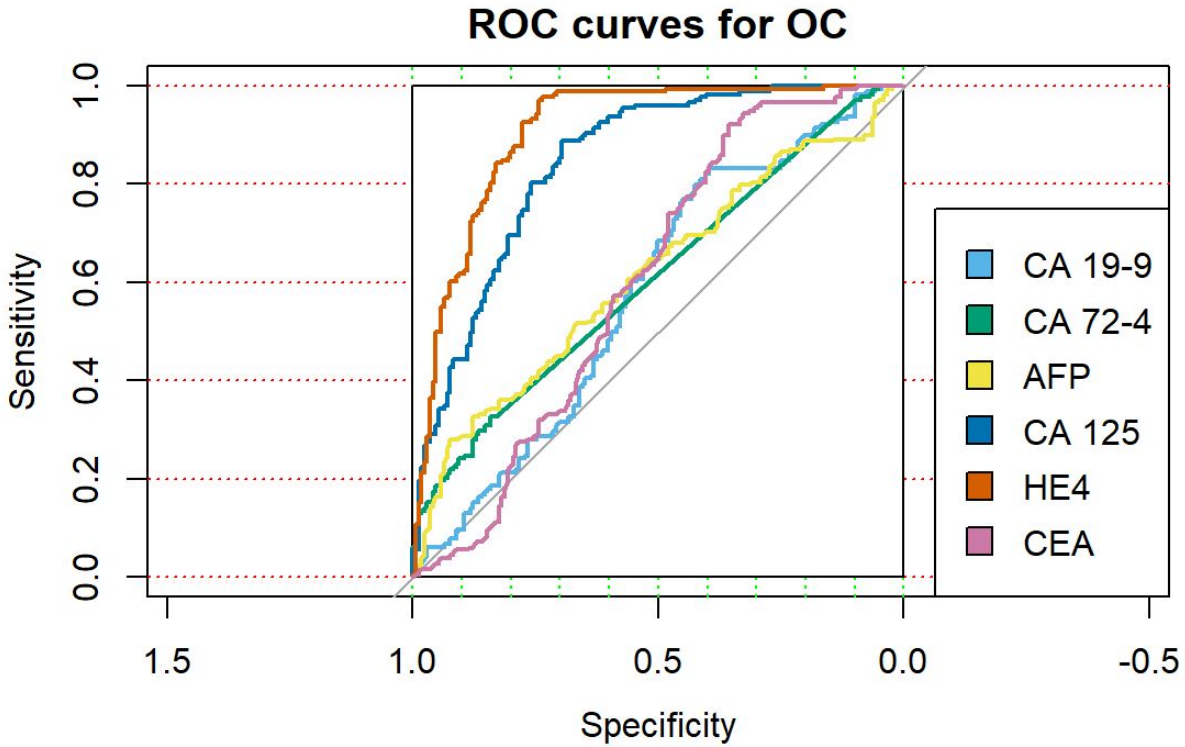
- Common: 1/75 people who have ovary
- Serious: undetected until late stage, fatal
- Diagnosis:
- Imaging - expensive
- Biopsy - very expensive, invasive
- Blood test - relatively affordable

Data

- obtained from 349 patient with tumor
- CA19-9: most commonly used for ovarian cancer
- CA72-4: most commonly used for gastric cancer
- AFP: related to cancer of the liver, ovaries or testicle
- CA125: related to ovarian cancer
- HE4: significant increase of HE4 for epithelial ovarian cancer
- CEA: certain types of cancer can increase your CEA levels, but you can have high CEA without having cancer

doi: 10.3390/jpm12081211

ROCs

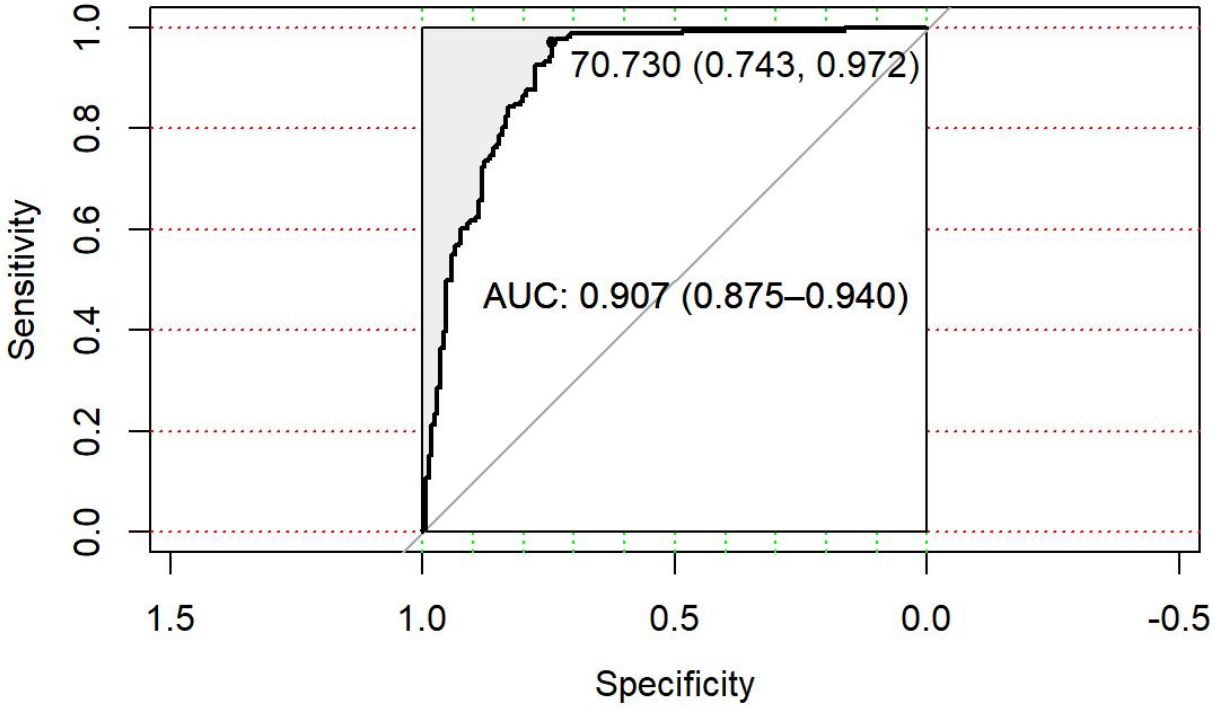

ROC curves for OC
Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1.5
1.0
0.5
0.0
-0.5
Specificity
CA 19-9
CA 72-4
AFP
CA 125
HE4
CEA

Compare: DeLong

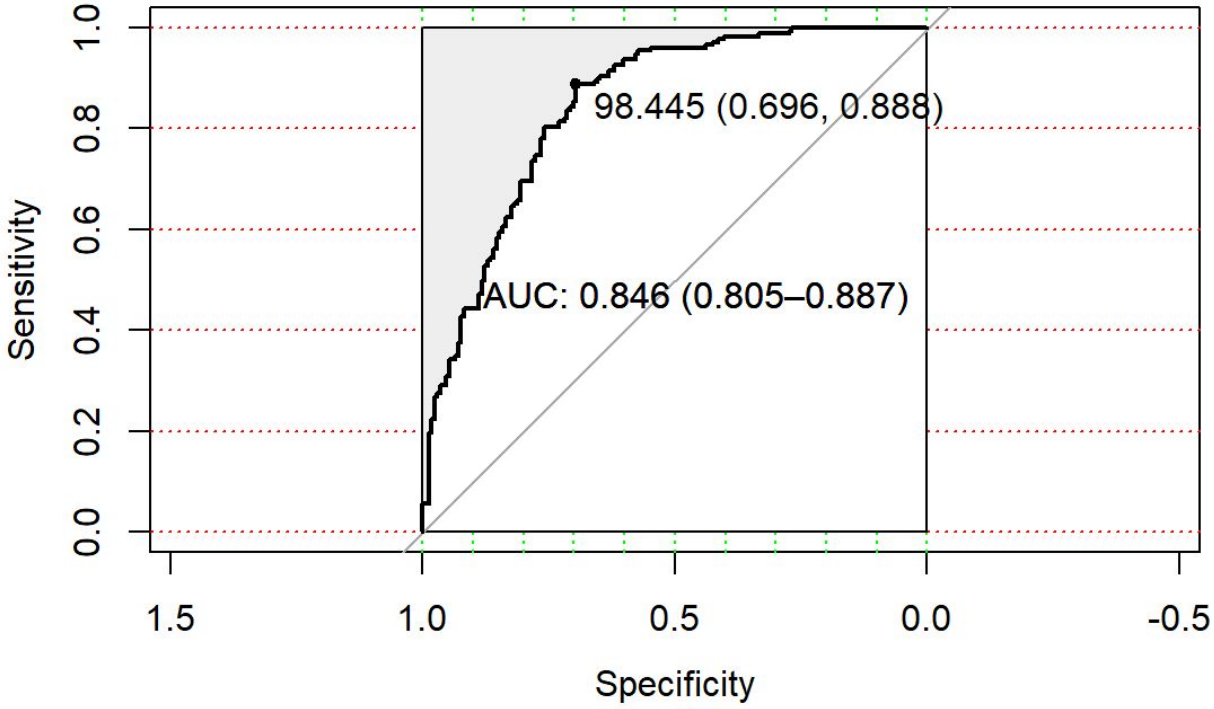
- The roc of HE4 has the highest AUC
- The roc of CA125 is the only one that has a relatively small difference with the roc of HE4

Compare	P-value
HE4 v.s. CA19-9	2.2e-16
HE4 v.s. CA72-4	2.2e-16
HE4 v.s. AFP	2.2e-16
HE4 v.s. CA125	0.002019
HE4 v.s. CEA	2.2e-16

ROC: HE4


70.730 (0.743, 0.972)
AUC: 0.907 (0.875–0.940)
Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1.5
1.0
0.5
0.0
-0.5
Specificity

ROC: CA125


98.445 (0.696, 0.888)
AUC: 0.846 (0.805–0.887)
Sensitivity
Specificity
0.0
0.2
0.4
0.6
0.8
1.0
1.5
1.0
0.5
0.0
-0.5

Downsides of HE4

- Expensive (400+ $ v.s. 200-$)
- Not available in most institutions

We choose HE4 and CA125!

HE4 cutoff

- Original: 70.73
- Modified: 71.44

- Other researchers from different countries found the same result of 70.

CA125 cutoff

- Original: 98.445
- Modified: 98.445

- Cutoff provided by manufacturer: 35

Determined from distribution in healthy individuals to include 99% of the normal population.
Shih M, et al.Tumor Markers: Physiology, Pathobiology, Technology and Clinical Applications. Washington, DC: AACC Press; 2002. 239-52.

Conclusion for this case

- HE4 is the best biomarker.
- For more affordable option, use CA125
- Extension:
- Study more data to get a cutoff with less error
- Explore if optimal biomarker & threshold depends on other individual characteristics such as age and menopause state

General Conclusion

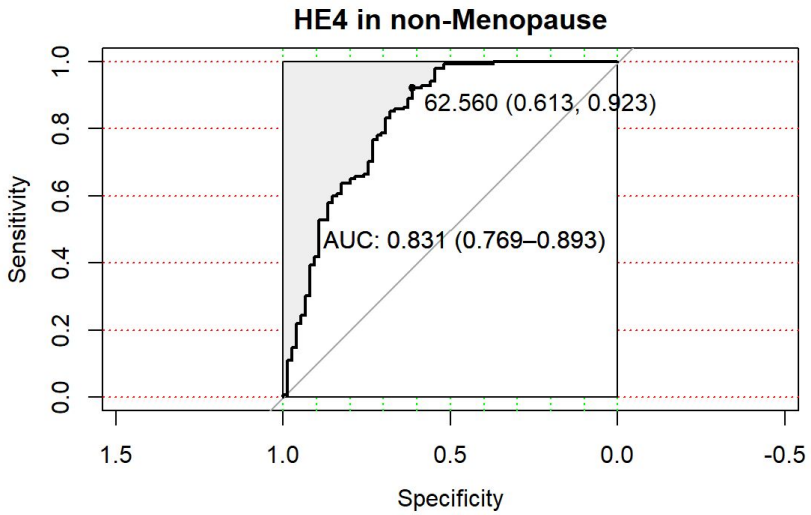
Conclusion

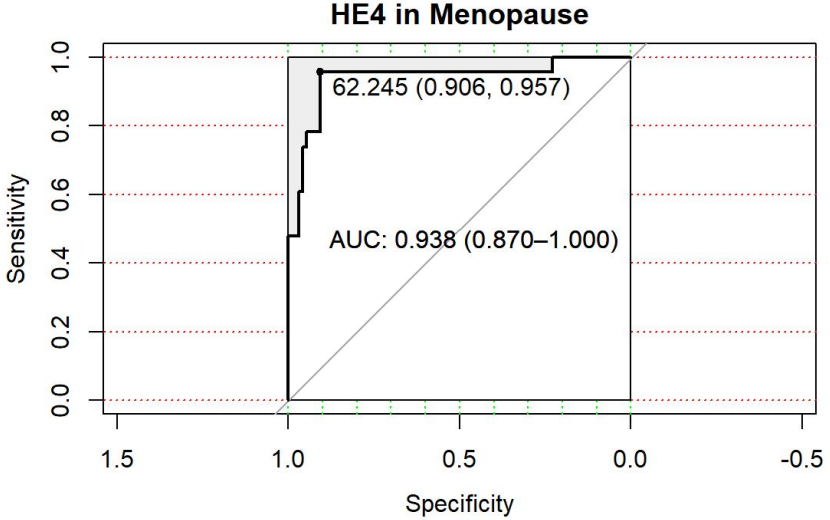
- ROC: Tool easy to understand
- Common, but there could be otherways
- Extension: combine biomarkers to a score, holistic view (e.g. age, sex)

Thank you

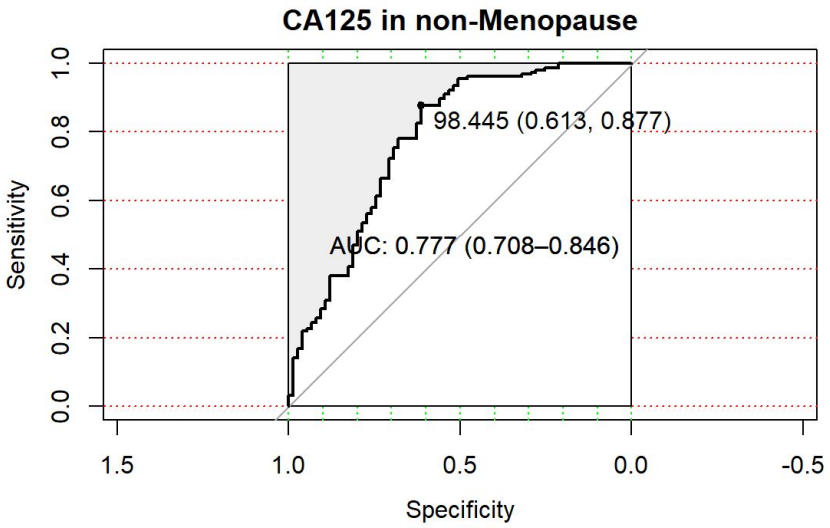
Further cutoff

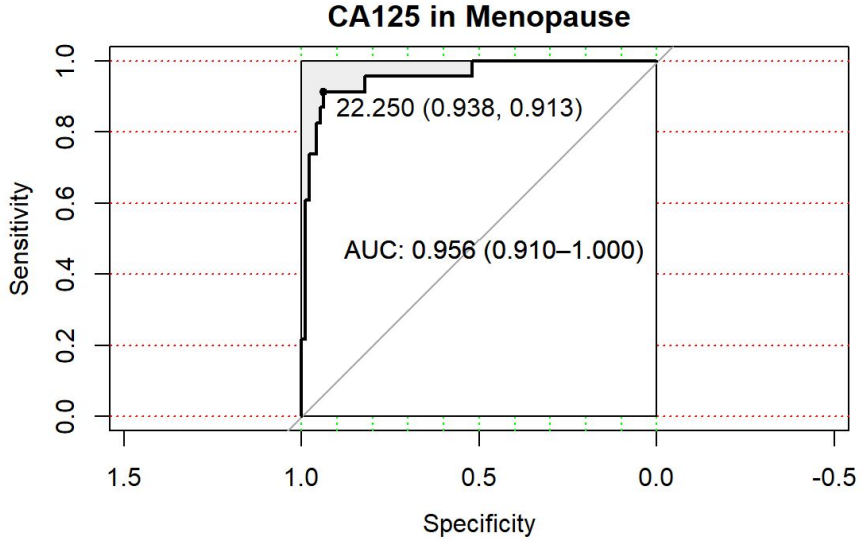
HE4


HE4 in non-Menopause
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
1.5
1.0
0.5
0.0
-0.5
Specificity
62.560 (0.613, 0.923)
AUC: 0.831 (0.769–0.893)


HE4 in Menopause
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
1.5
1.0
0.5
0.0
-0.5
Specificity
62.245 (0.906, 0.957)
AUC: 0.938 (0.870–1.000)

CA125


CA125 in non-Menopause
98.445 (0.613, 0.877)
AUC: 0.777 (0.708–0.846)
Sensitivity
Specificity
0.0
0.2
0.4
0.6
0.8
1.0
1.5
1.0
0.5
0.0
-0.5


CA125 in Menopause
22.250 (0.938, 0.913)
AUC: 0.956 (0.910–1.000)
Sensitivity
Specificity
0.0
0.2
0.4
0.6
0.8
1.0
1.5
1.0
0.5
0.0
-0.5