

Alternative Metrics for Virtual Screening

Oleksandr Voroshylov

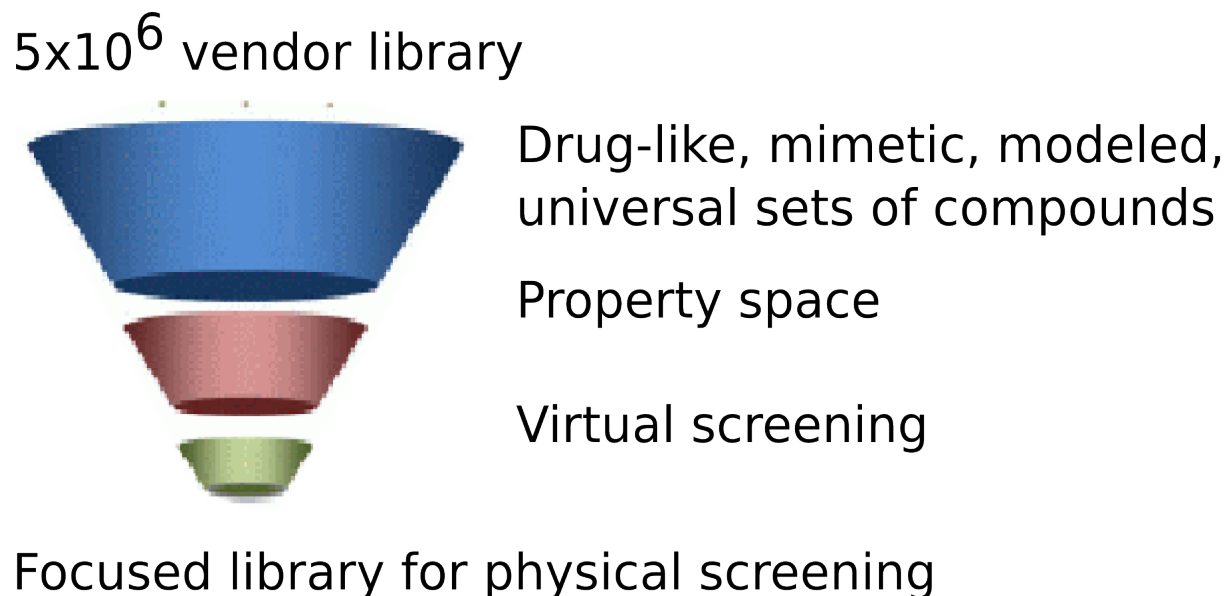
Nicholls A (2008) What do we know and when do we know it?
J Comput Aided Mol Des 22:239–255

Truchon J-F, Bayly CI (2007) Evaluating virtual screening methods: good and bad metrics for the “early recognition” problem.
J Chem Inf Model 47:488–508

Agenda

- Virtual screening
- Existing quality measures
- Problems of measurement metrics
- Cost structure of virtual screening

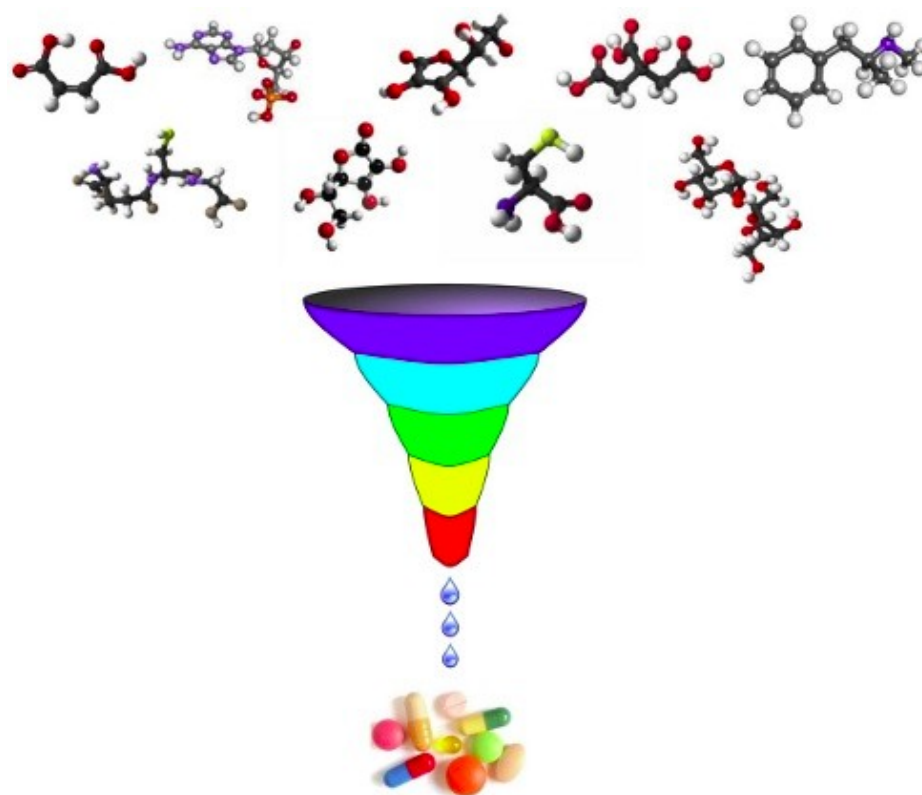
What is Virtual Screening?



Computational technique to search libraries of small molecules in order to identify structures which are most likely to bind to a drug target

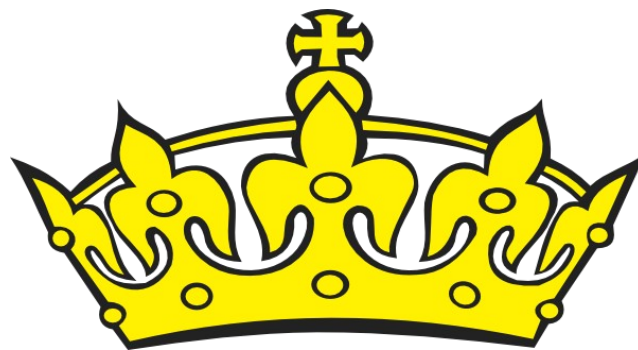
What is Virtual Screening?

- Library of small molecules
- Protein receptor or enzyme
- Computational technique
- Evaluation metrics



What is expected from Virtual Screening?

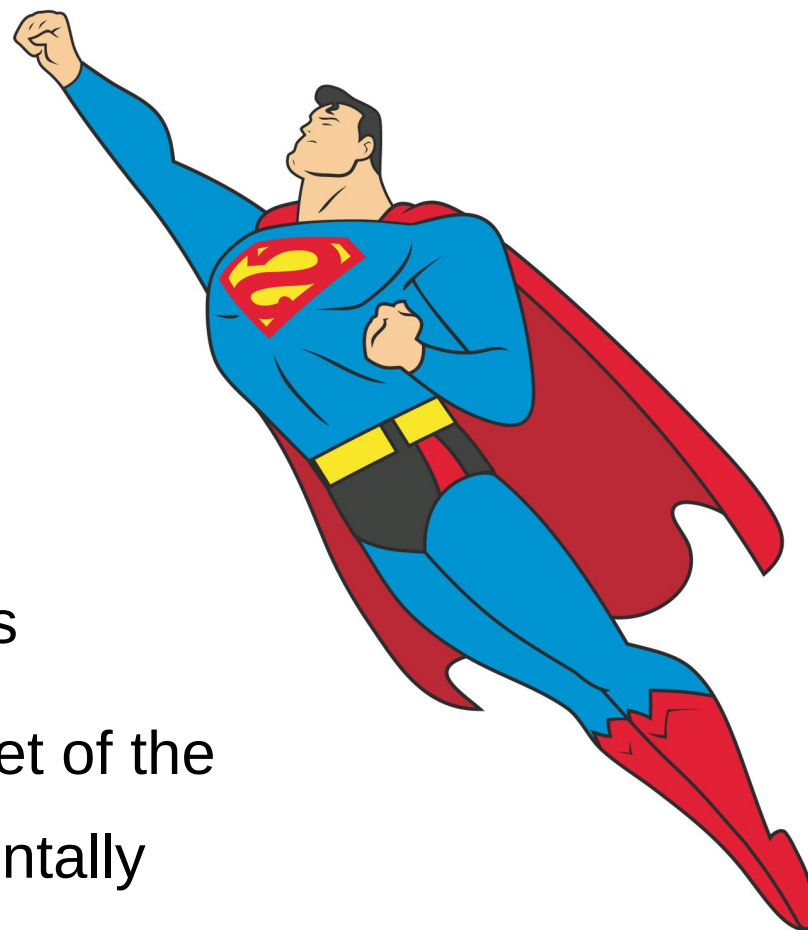
1. Approach works
2. Very small set of compounds will be tested experimentally



What is expected from Virtual Screening?

Early recognition

- Ranks actives very early
- Works with larger set of compounds
- Makes it possible that very small set of the compounds will be tested experimentally



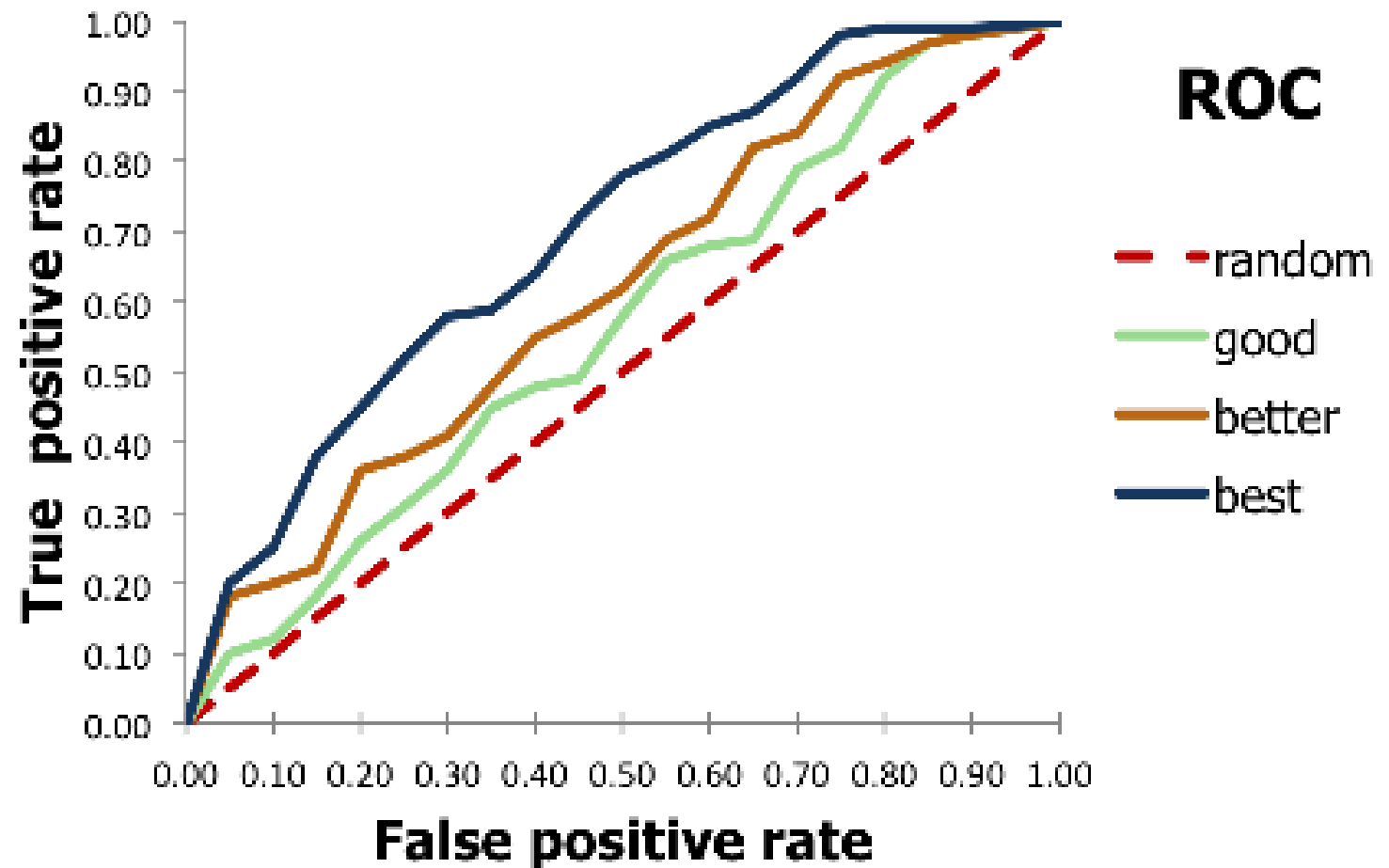
What is expected from Virtual Screening?

What about point 1 (Approach works) ?

Traditional Measures

- Area under Receiver Operating Characteristic
- Area under Accumulation Curve
- Enrichment Factor

Receiver operating characteristic (ROC)



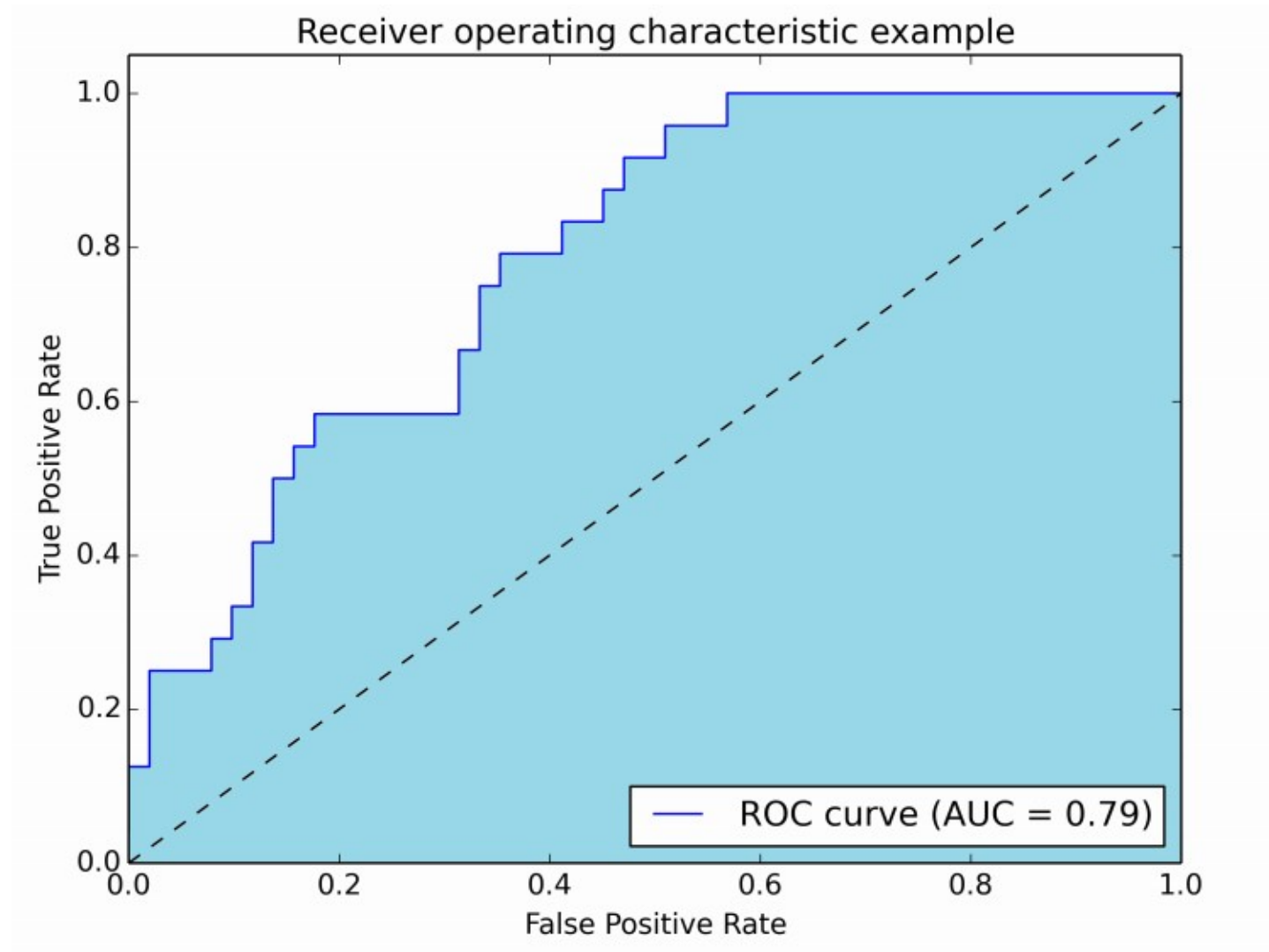
Receiver operating characteristic (ROC)

- **TP** – true positive
- **TN** – true negative
- **FP** – false positive
- **FN** – false negative

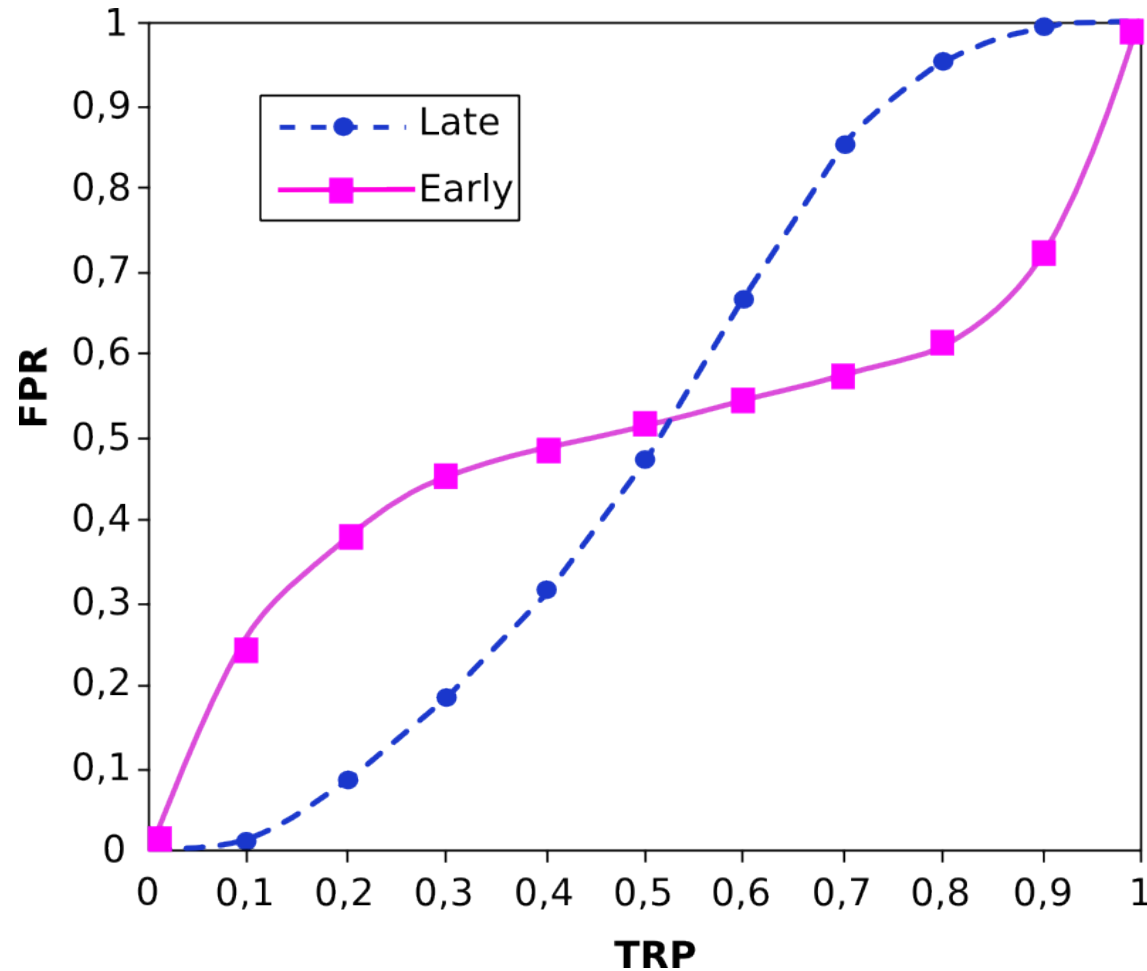
$$TPR = TP / (TP + FN)$$

$$FPR = FP / (FP + TN)$$

Area under ROC



Why is AUC not enough?



Area under ROC = 0.5

Alternatives

- Robust Initial Enhancement (RIE)
 - developed by Sheridan
- Boltzmann-Enhanced Discrimination of ROC (BEDROC)
 - developed by Truchon

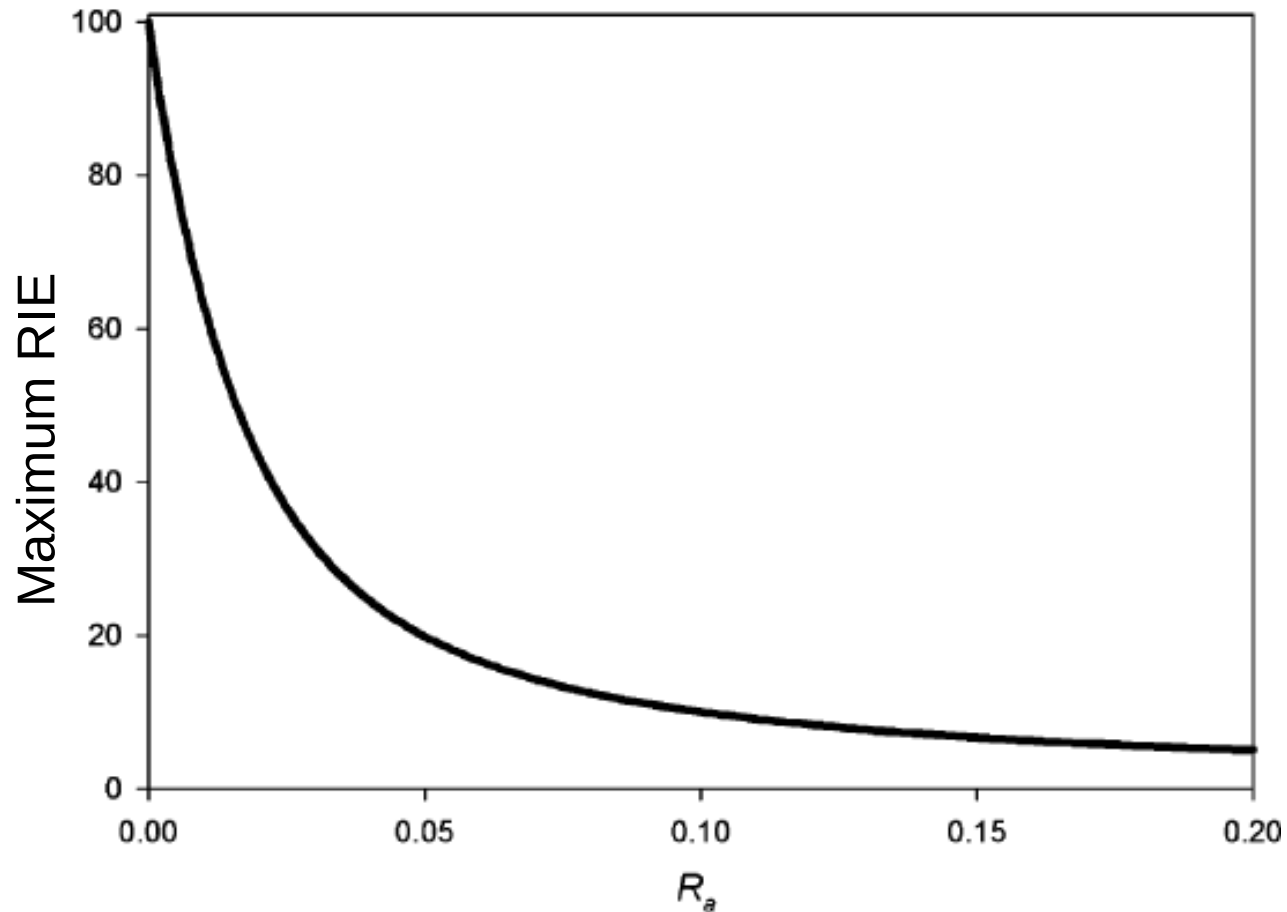
Robust Initial Enhancement (RIE)

$$RIE = \frac{\sum_{i=1}^n e^{\frac{-\alpha r_i}{N}}}{\left\langle \sum_{i=1}^n e^{\frac{-\alpha r_i}{N}} \right\rangle_r}$$

Metric using a continuously decreasing exponential weight as a function of rank

- r_i - is a rank of the i -th active
- N - is a number of elements
- n - is a number of actives
- α - is a exponential factor
- angle brackets means averaging
- subscript r means that it is over a uniform distribution.

Robust Initial Enhancement (RIE)



RIE maximum as a function of the ratio of actives R_a with $a=100$.

Boltzmann-Enhanced Discrimination of ROC (BedROC)

$$BEDROC = \frac{RIE - RIE_{min}}{RIE_{max} - RIE_{min}}$$

Values from 0 to 1

RIE and BedROC Properties

RIE

- Gives early rankings more weight
- Weight depends on their position in the list
- Exponential factor
- Normalized by RIE of a random distribution of actives
- Sensitive to the total number of inactives

BedROC

- Gives early rankings more weight
- Weight depends on their position in the list
- Exponential factor
- Normalization by the maximum dynamic range
- Sensitive to the total number of inactives
- Values from 0.0 to 1.0

Are they perfect?

RIE

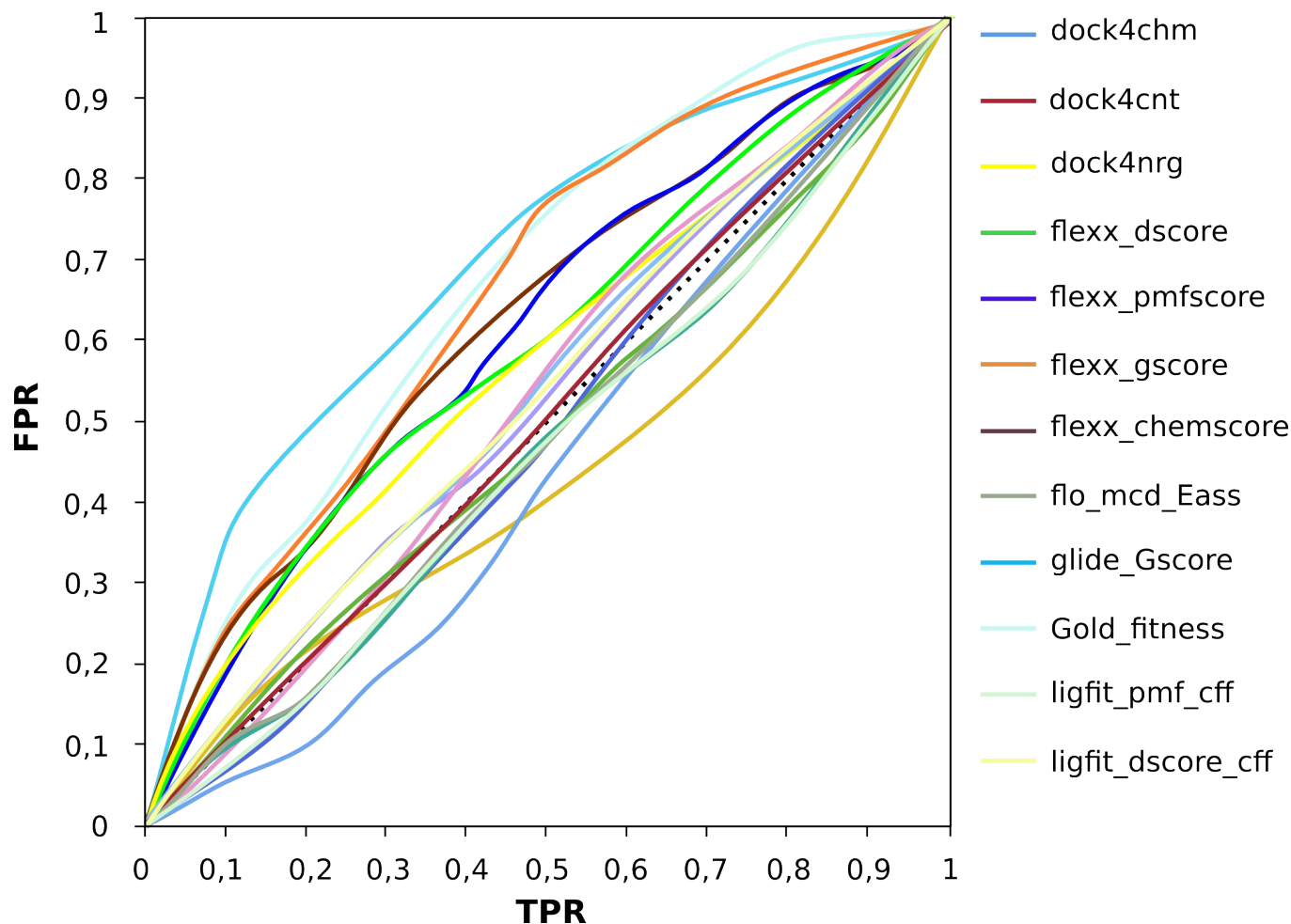
- Have RIE values a absolute meaning?
- Are RIE values interpretable?
- How good is the RIE value 5.35?

BedROC

- Have BedROC values an absolute meaning?
- BedROC scales the values from 0 to 1, how understandable is that?
- Can you compare BedROC values with different alpha-factors ?

Are they better as AUC?

**Twenty ROC Curves Averaged over Eight Targets.
From the Warren et al/ GSK Dataset**

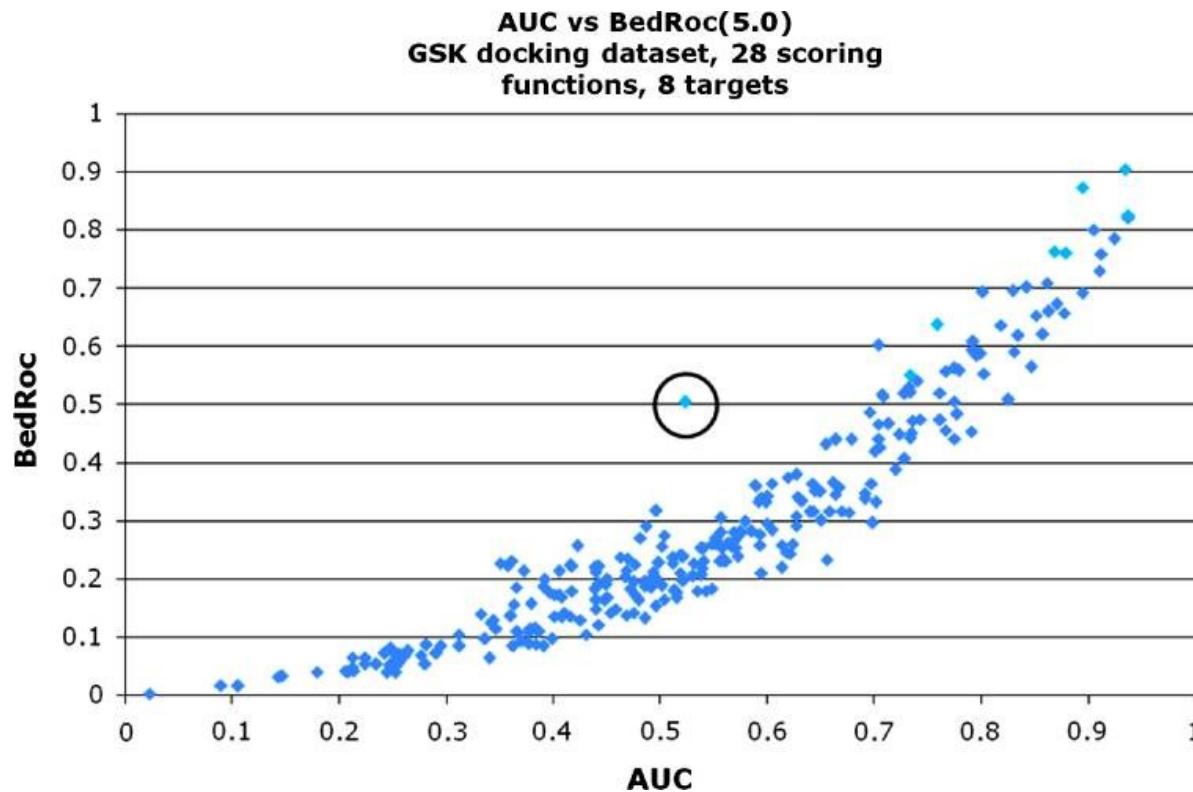


No biphasic nature.

Why there is no biphasic nature?

- The individual curves are not biphasic.
- The averaging dilutes this characteristic.

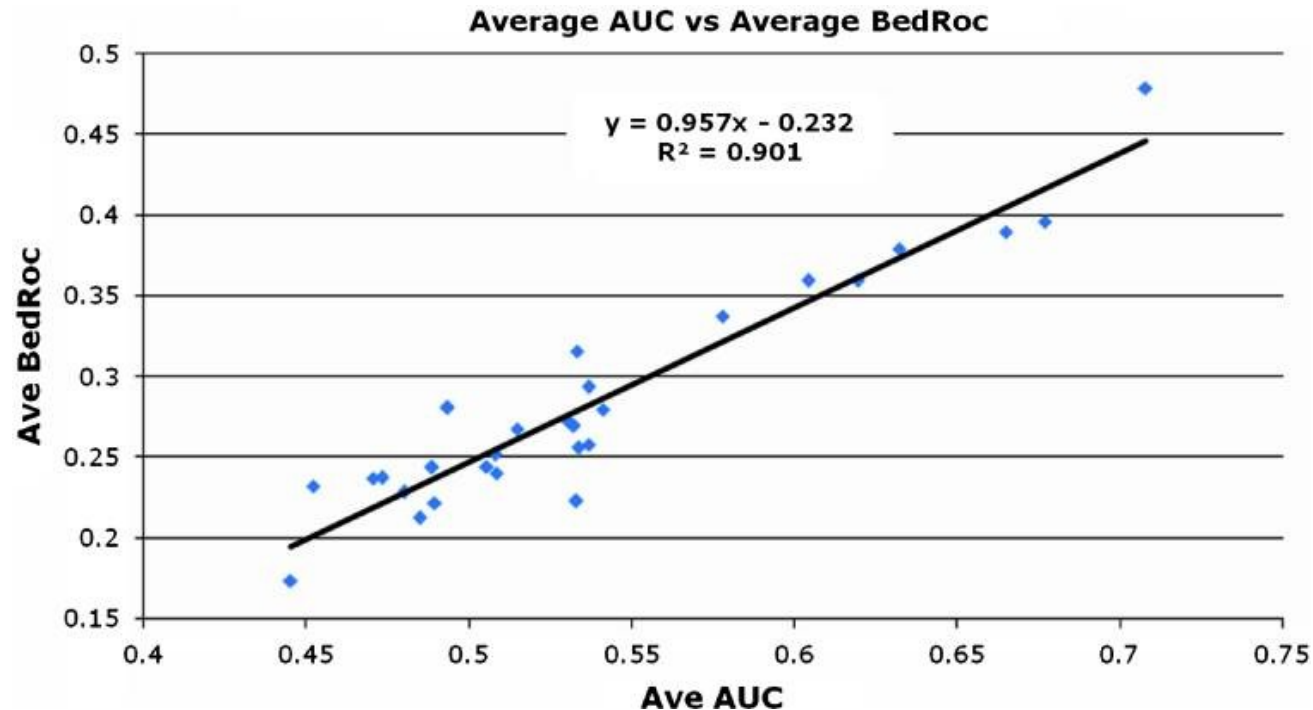
270 Virtual Screenings



- Low AUC
- High BedROC

No evidence for biphasic behavior

Average biphasic behavior?



The better the AUC value is, the better the BedROC value is (strongly correlated).

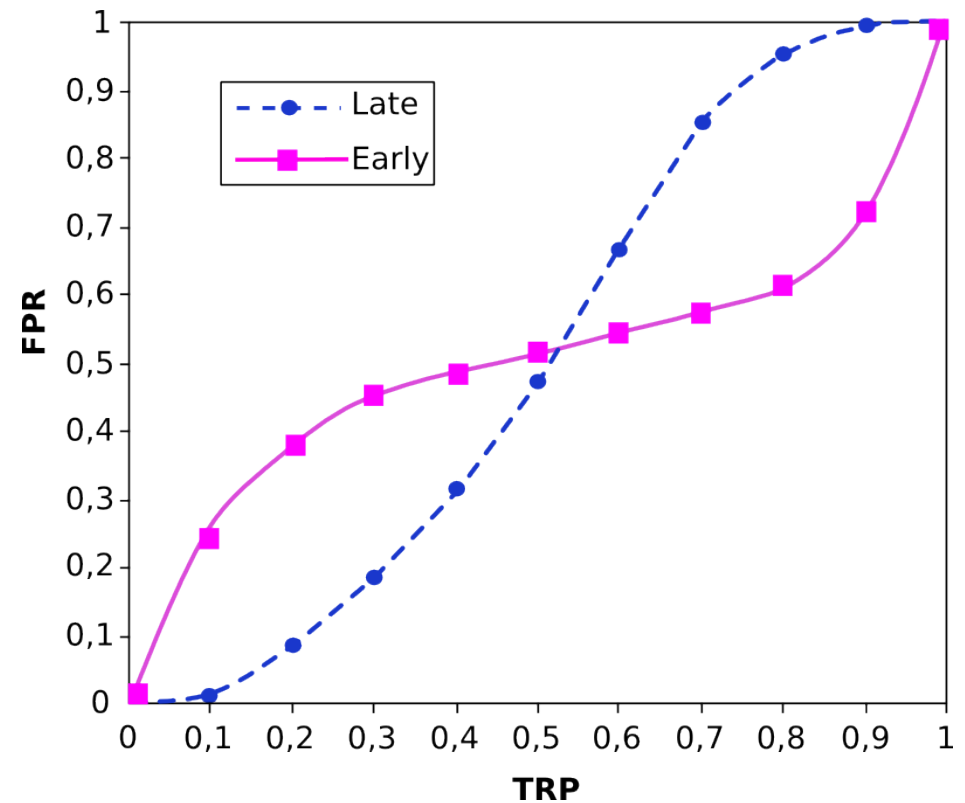
What does this mean?

- No evidence for biphasic nature
- The better the AUC value is, the better the BedROC value is
- In average case RIE and BedROC are not better than AUC

Is there something better?

- Practice-orientated
- Easy to understand

Gives an answer why the solid curve represents a better method than the dashed curve



Is there something better?



Cost structure of Virtual Screening

- Easy to understand
 - also for outsider
- Industry oriented

What is a cost structure of Virtual Screening?

True positive values = successful diagnosis

- Fast results
- Saves time
- Saves money

What is a cost structure of Virtual Screening?

False positive values = false diagnosis

- Costly tests
- Waste of time
- Waste of money

What is a cost structure of Virtual Screening?

False negative values = more severe condition develops

- May cost a lot
- Waste of time
- Waste of money

What is a cost structure of Virtual Screening?

True negative values – no false diagnosis

- Saves money
- Saves time

What is a cost structure of Virtual Screening?

To transform ROC into real costs we need:

- Expected number of actives and inactives
- Or the ratio of the two

Cost structure of Virtual Screening

- True positives = 8.0
- True negatives = 0.02
- False positives = -0.16
- False negatives = -2.0

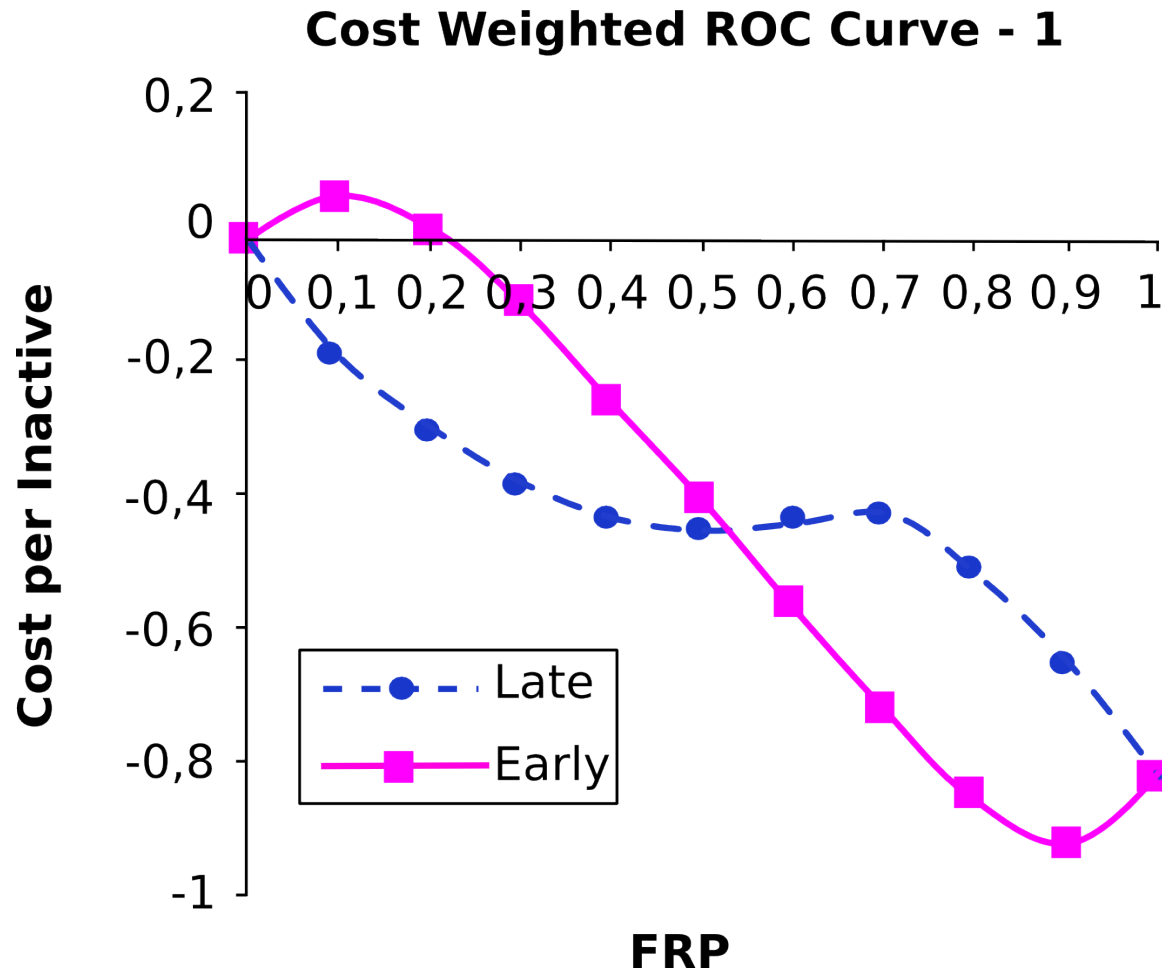
$$\begin{aligned} Cost(t) = & TPR \times N_a \times (8.0) + (1 - TPR) \times \\ & N_a \times (-2.0) + FPR \times N_i \times (-0.16) + \\ & (1 - FPR) \times N_i \times (0.02) \end{aligned}$$

Cost structure of Virtual Screening

$$N_a/N_i = 1/100$$

$$Cost(t)/N_i = 0.10 \times TPR - 0.18 \times FPR$$

Cost weighted versions of ROC



- TP = 8.0
- FN = -2.0
- FP = -0.16
- TN = 0.02

$$Cost(t)/N_i = 0.1 \times TPR - 0.18 \times FPR$$

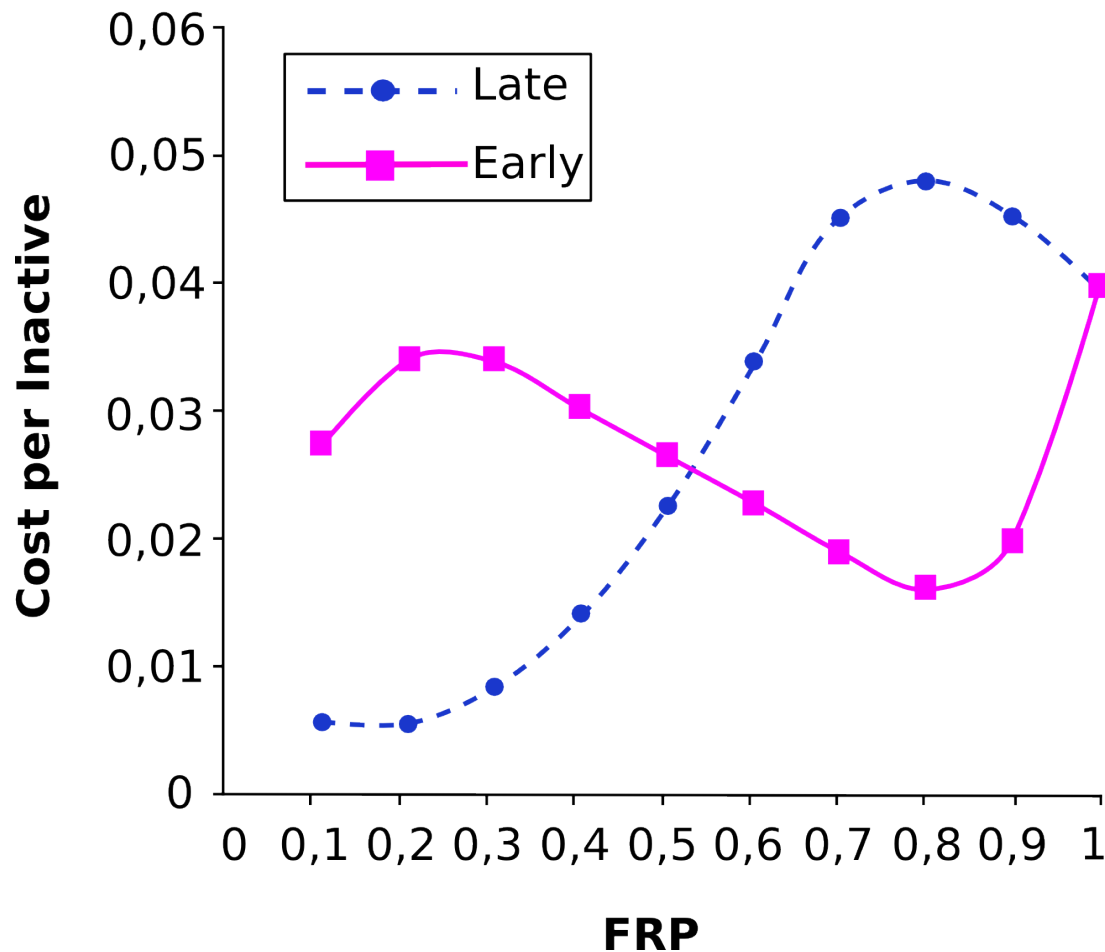
What does this mean?

Early recognition is better as late recognition

Late recognition ist never cost effective

Reducing the cost of a FP by 75%

Cost Weighted ROC Curve - 2



- TP = 8.0
- FN = -2.0
- FP = -0.04
- TN = 0.03

$$Cost(t)/N_i = 0.1 \times TPR - 0.07 \times FPR + 0.01$$

What does this mean?

By reducing the cost of a false positive by 75%
late recognition is also cost effective.

Assumptions

- These examples are obviously only illustrative.
- Early recognition is important only because of an assumed cost structure.

Summary

- Virtual screening is an important part of a drug-discovery
- Metrics for Virtual Screening are the common problem
- Cost structure of Virtual Screening might be a solution
- This is an open field for research