

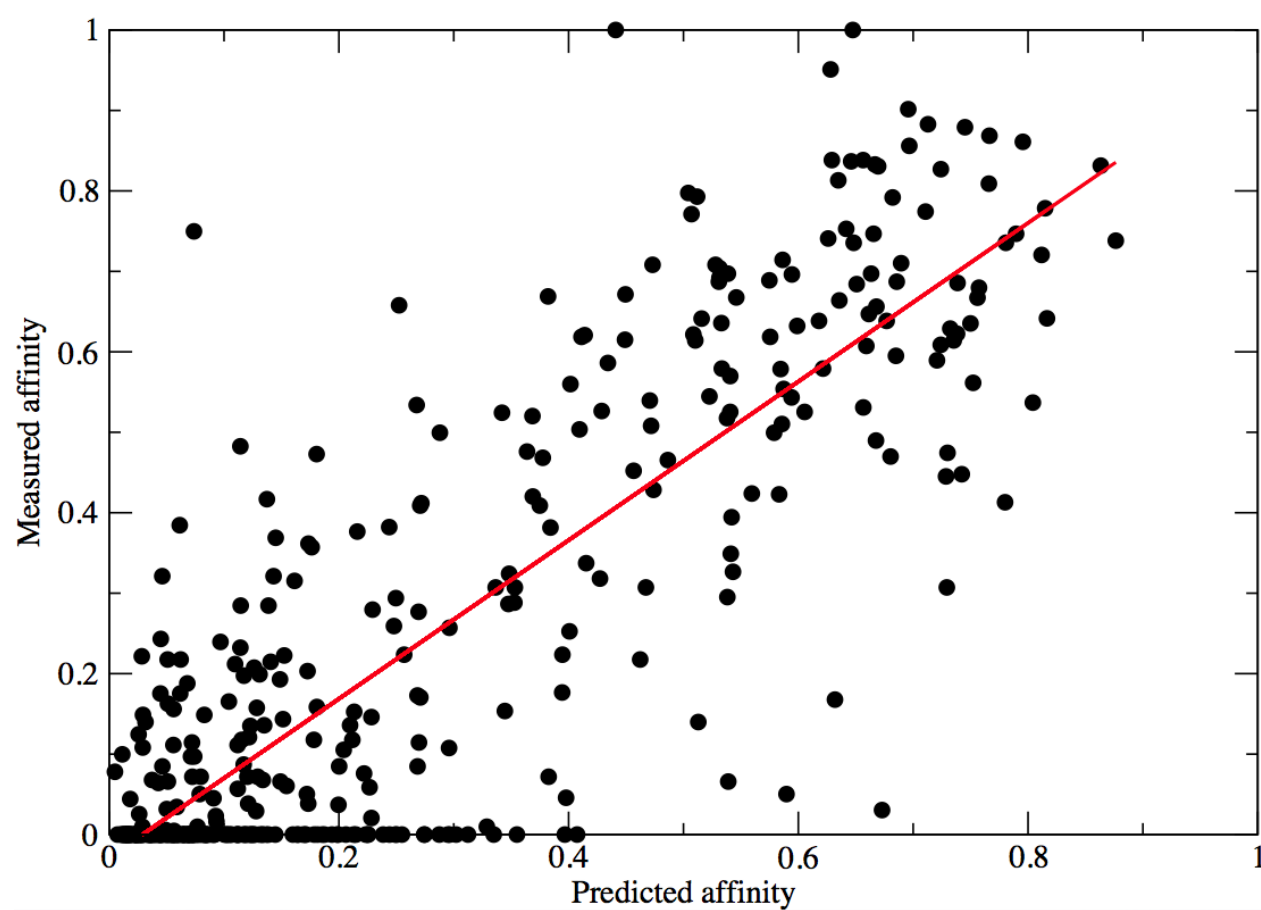
# Performance measures

Morten Nielsen,  
CBS, Department of Systems Biology,  
DTU

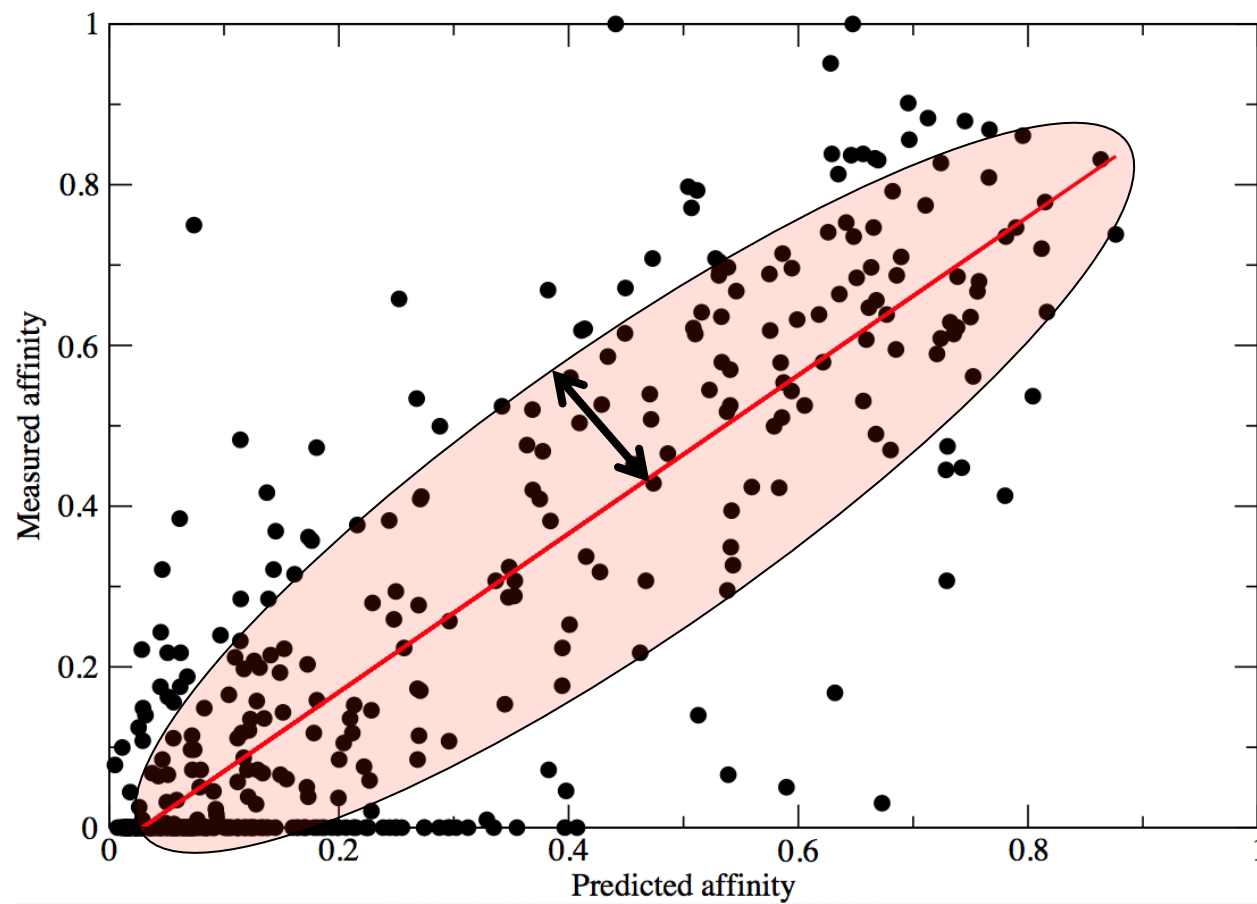
---

# Performance measures

Meas	Pred
0.4050	0.8344
0.9373	1.0000
0.8161	0.6388
0.6752	0.9841
0.0253	0.0000
0.3196	0.5388
0.6764	0.6247
0.1872	0.1921
0.4220	0.6546
0.6545	0.6546
0.7917	0.1342
0.4405	0.3551
0.1548	0.0000
0.2740	0.1993
0.4399	0.6461
0.1725	0.3916
0.0539	0.0000
0.3795	0.5623
0.2242	0.1968
0.3108	0.2114
0.2260	0.0336
0.2780	0.5647
0.0198	0.1224
0.5890	0.5538
0.5120	0.4349
0.7266	1.0000
0.1136	0.0000
0.0456	0.2128
0.0069	0.4100
0.4502	0.3848

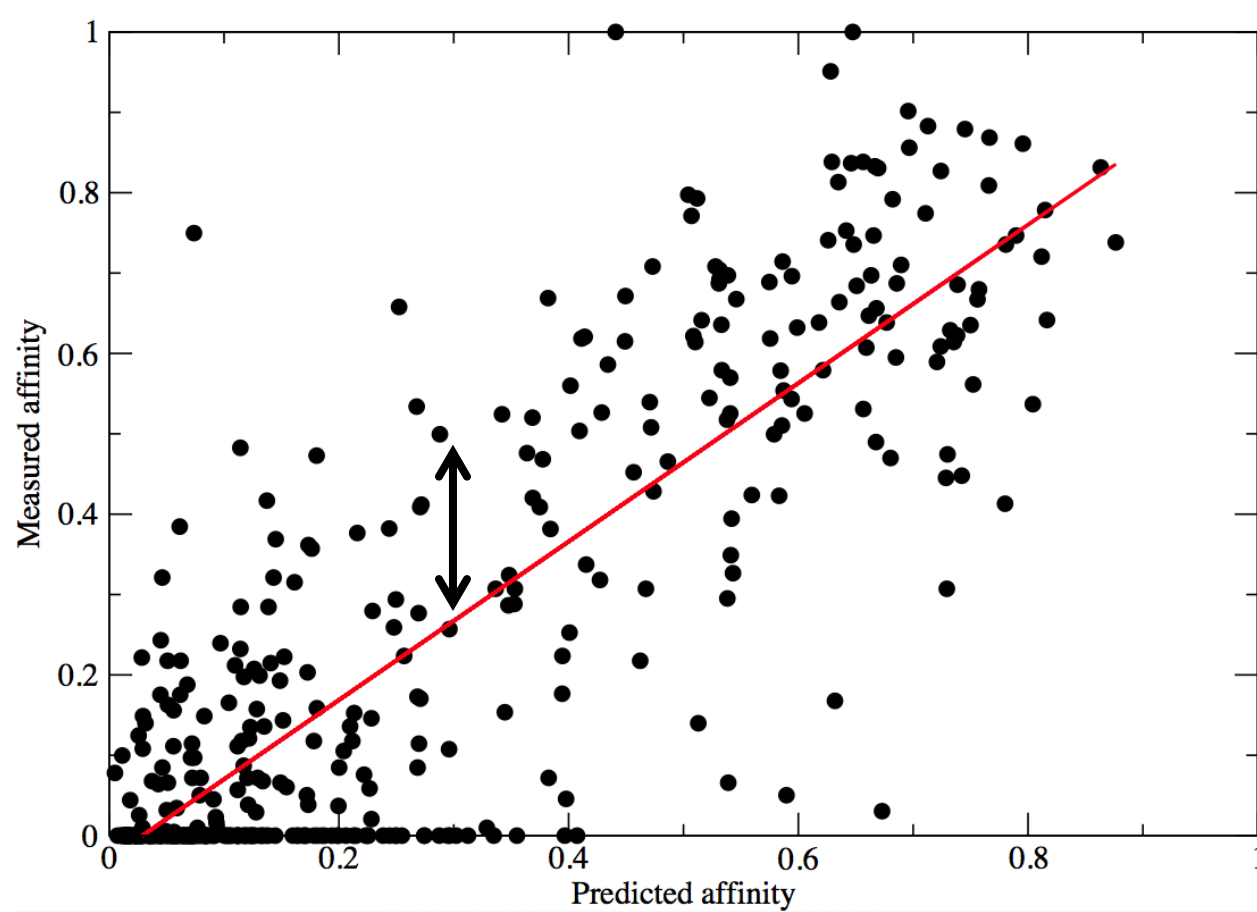


# Pearsons' correlation coefficient



$$PCC = \frac{\sum_i (a_i - \bar{a}) \cdot (p_i - \bar{p})}{\sqrt{\sum_i (a_i - \bar{a})^2} \cdot \sqrt{\sum_i (p_i - \bar{p})^2}}$$

# Mean Squared Error

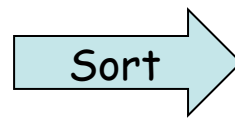


$$MSE = \frac{1}{N} \sum_i (a_i - p_i)^2$$

# Quantitative data

- Accuracy of prediction method

Prediction score	Binder/non-binder
0,94	1
0,88	1
0,12	0
0,43	1
0,58	0
0,66	1
0,77	0
0,11	0
0,77	0
0,06	1
0,71	1
0,43	0
0,05	0
0,79	1
0,55	0
0,32	1
0,22	1
0,33	0
0,35	0
0,57	1
0,21	0
0,91	1
0,16	0
0,45	1
0,88	0
0,66	1
0,98	1



Prediction score	Binder/non-binder
0,98	1
0,94	1
0,91	1
0,88	1
0,88	0
0,79	1
0,77	0
0,77	0
0,71	1
0,66	1
0,66	1
0,58	0
0,57	1
0,55	0
0,45	1
0,43	1
0,43	0
0,35	0
0,33	0
0,32	1
0,22	1
0,21	0
0,16	0
0,12	0
0,11	0
0,06	1
0,05	0

# Matthews correlation - Threshold of 0.5

Prediction score	Binder/non-binder
0,98	1
0,94	1
0,91	1
0,88	1
0,88	0
0,79	1
0,77	0
0,77	0
0,71	1
0,66	1
0,66	1
0,58	0
0,57	1
0,55	0
0,45	1
0,43	1
0,43	0
0,35	0
0,33	0
0,32	1
0,22	1
0,21	0
0,16	0
0,12	0
0,11	0
0,06	1
0,05	0

False negative	True positive
True negative	False positive

$$\text{Sensitivity} = TP / AP$$

$$\text{Specificity} = TN / AN$$

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FN)(TN + FP)(TP + FP)(TN + FN)}}$$

# Matthews correlation - Threshold of 0.5

Prediction score	Binder/non-binder
0,98	1
0,94	1
0,91	1
0,88	1
0,88	0
0,79	1
0,77	0
0,77	0
0,71	1
0,66	1
0,66	1
0,58	0
0,57	1
0,55	0
0,45	1
0,43	1
0,43	0
0,35	0
0,33	0
0,32	1
0,22	1
0,21	0
0,16	0
0,12	0
0,11	0
0,06	1
0,05	0

5	9
8	5

$$MCC = \frac{9 \cdot 8 - 5 \cdot 5}{\sqrt{(9 + 5)(8 + 5)(9 + 5)(8 + 5)}} = 0.258$$

# Performance measure - Roc curve

Prediction score	Binder/non-binder
0,98	1
0,94	1
0,91	1
0,88	1
0,88	0
0,79	1
0,77	0
0,77	0
0,71	1
0,66	1
0,66	1
0,58	0
0,57	1
0,55	0
0,45	1
0,43	1
0,43	0
0,35	0
0,33	0
0,32	1
0,22	1
0,21	0
0,16	0
0,12	0
0,11	0
0,06	1
0,05	0

False negative	True positive
True negative	False positive

Threshold	TP	FN	TP/(TP+FN)	FP	TN	FP/(FP+TN)
>0,8	4	10	0.29	1	12	0.08
>0,6						
>0,4						
>0,2						
>0						



# Performance measure - Roc curve

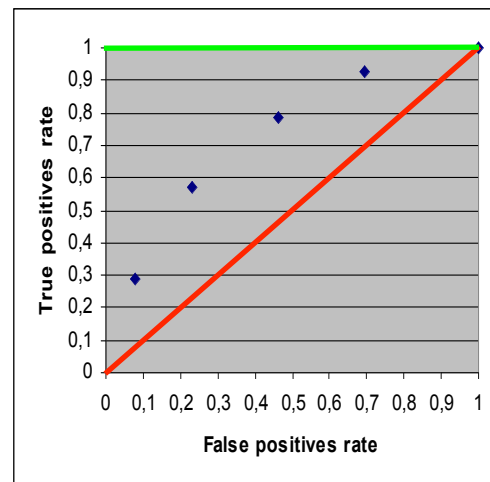
Prediction score	Binder/non-binder
0,98	1
0,94	1
0,91	1
0,88	1
0,88	0
0,79	1
0,77	0
0,77	0
0,71	1
0,66	1
0,66	1
0,58	0
0,57	1
0,55	0
0,45	1
0,43	1
0,43	0
0,35	0
0,33	0
0,32	1
0,22	1
0,21	0
0,16	0
0,12	0
0,11	0
0,06	1
0,05	0

False negative	True positive
True negative	False positive

Threshold	TP	FN	TP/(TP+FN)	FP	TN	FP/(FP+TN)
>0,8	4	10	0.29	1	12	0.08
>0,6	8	6	0.57	3	10	0.23
>0,4						
>0,2						
>0						

# ROC curves

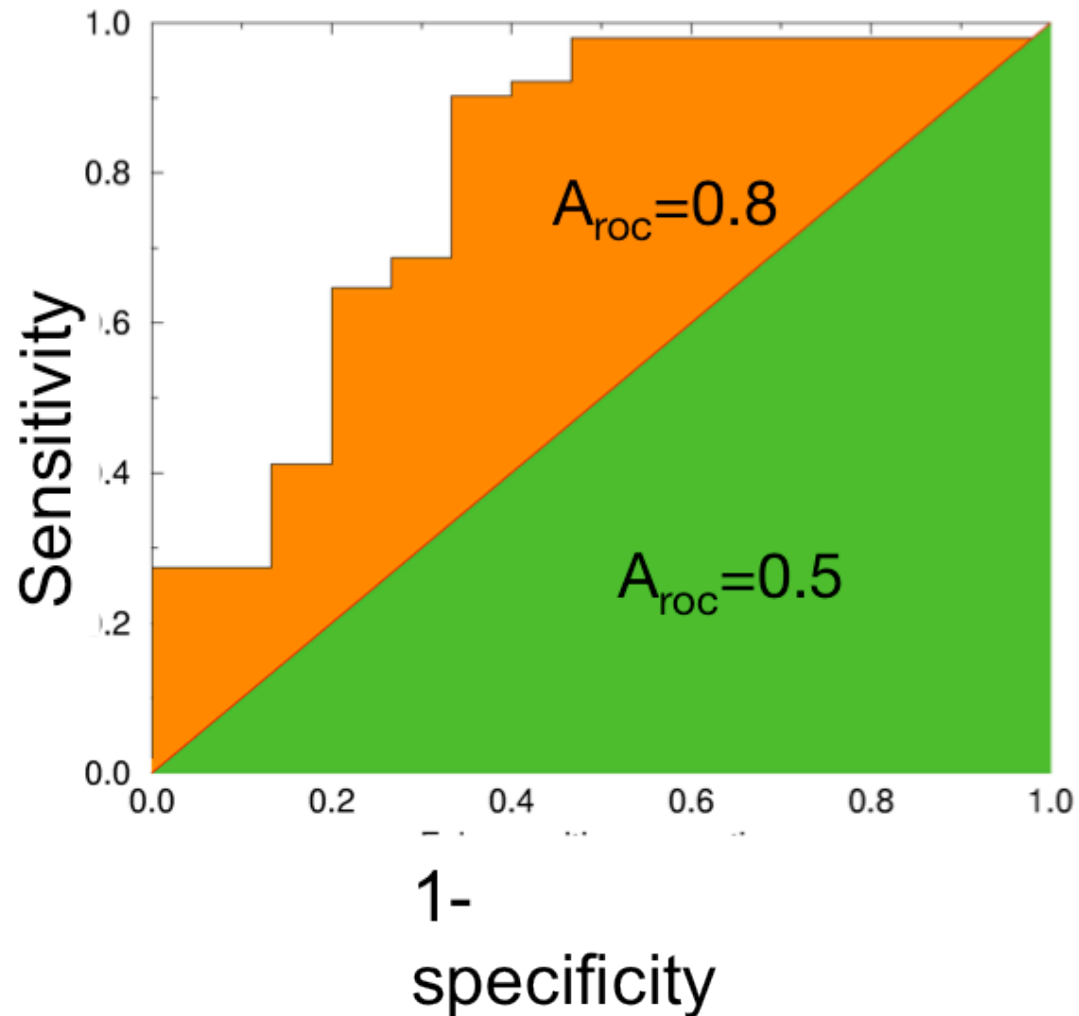
Threshold	TP	FN	TP/(TP+FN)	FP	TN	FP/(FP+TN)
>0.8	4	10	0.29	1	12	0.08
>0.6	8	6	0.57	3	10	0.23
>0.4	11	3	0.79	6	7	0.46
>0.2	13	1	0.93	9	4	0.69
>0	14	0	1	13	0	1



**AUC = 0.5**  
**AUC = 1.0**

# AUC (area under the ROC curve)

Roc curves



# Summary

---

- **MCC and PCC**
    - Random = 0.0
    - Perfect = 1.0 (or -1)
  - **ROC (AUC)**
    - Random = 0.5
    - Perfect = 1 (or 0)
-