True positive rate (Sensitivity)

$$\text{true positive rate} = \frac{\text{\# of true positives}}{\text{\# of known positives}}$$

(Proportion of actual positives that are correctly identified)

True negative rate (Specificity)

$$\text{true negative rate} = \frac{\text{\# of true negatives}}{\text{\# of known negatives}}$$

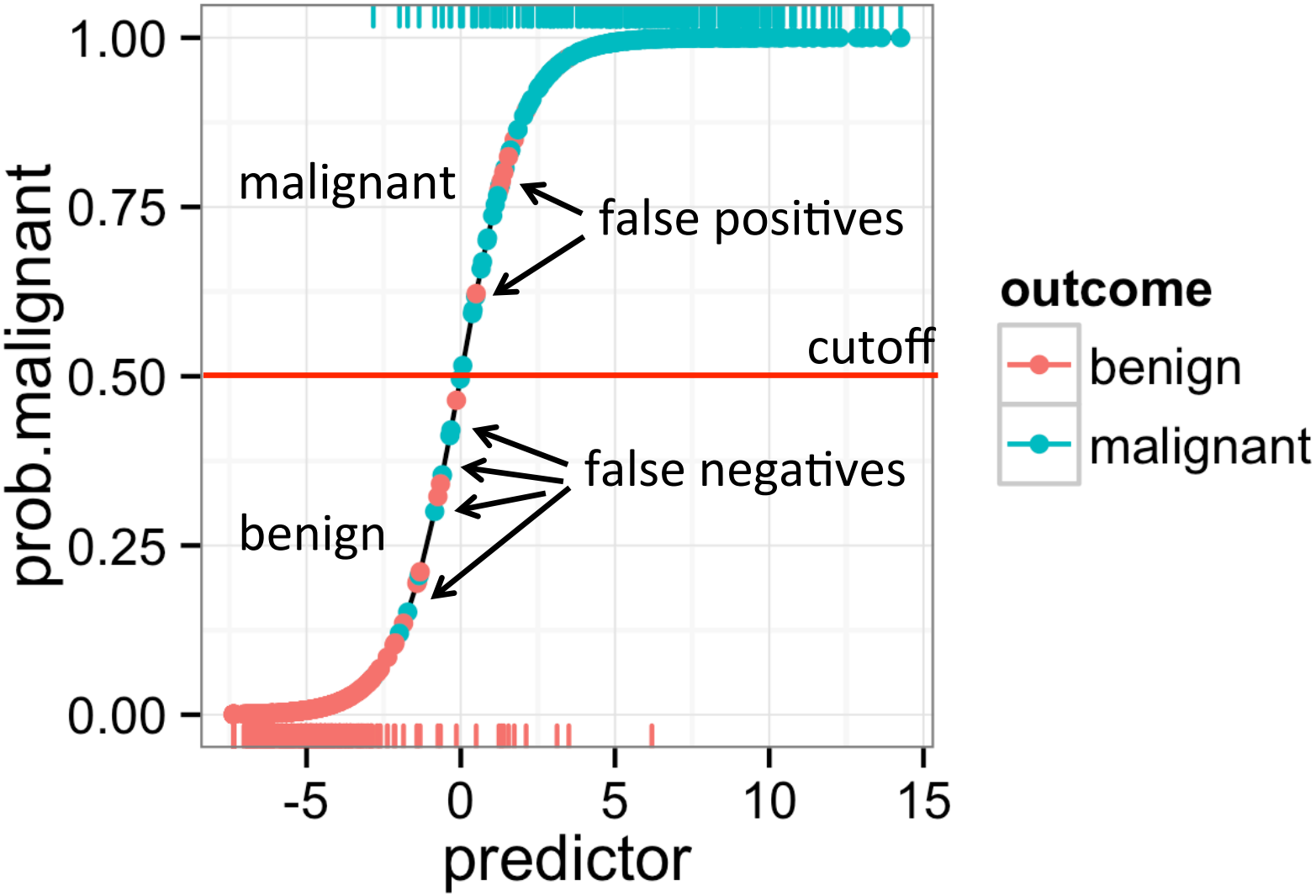
(Proportion of actual negatives that are correctly identified)

False positive rate (1 – Specificity)

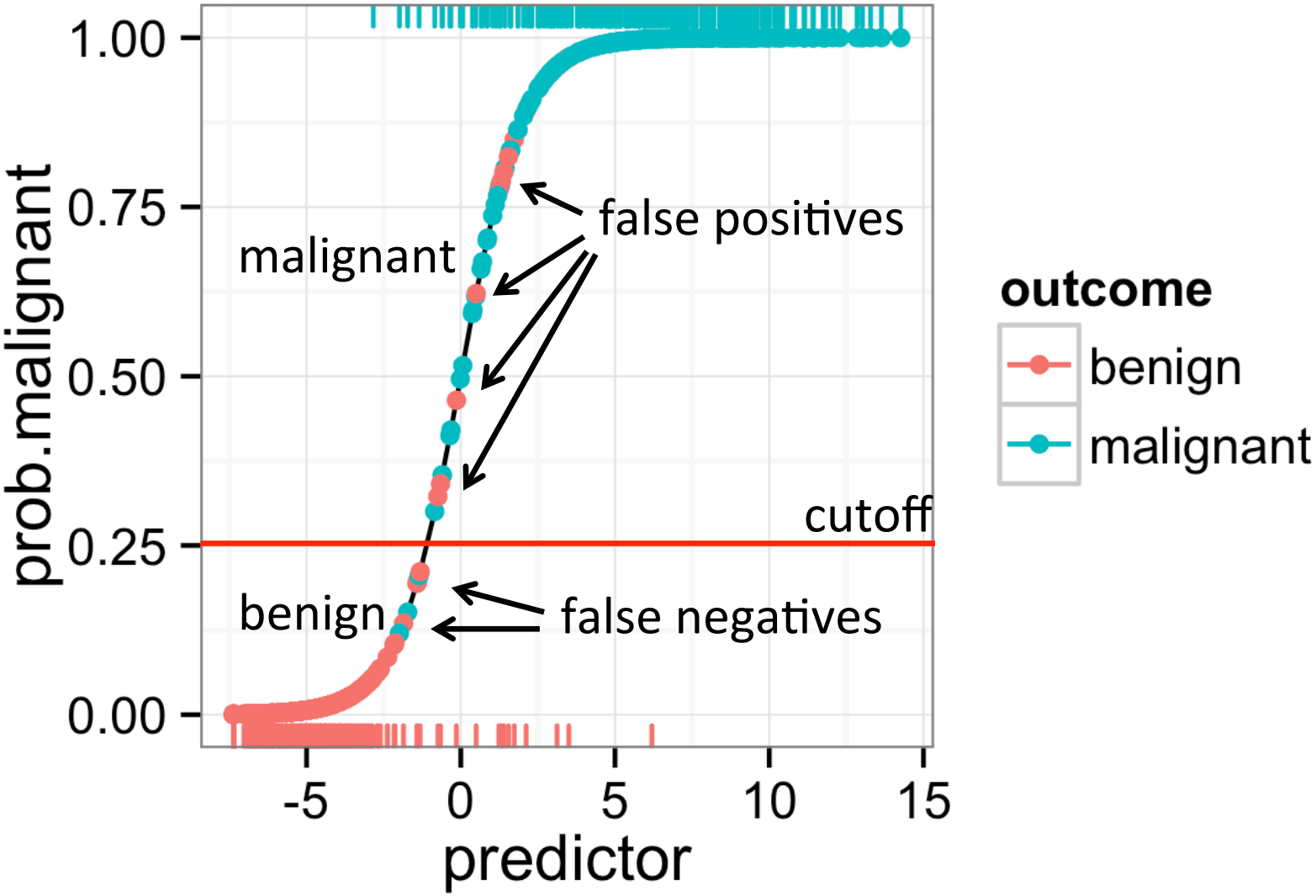
$$\text{false positive rate} = \frac{\text{\# of false positives}}{\text{\# of known negatives}}$$

(Proportion of actual negatives that are incorrectly identified)

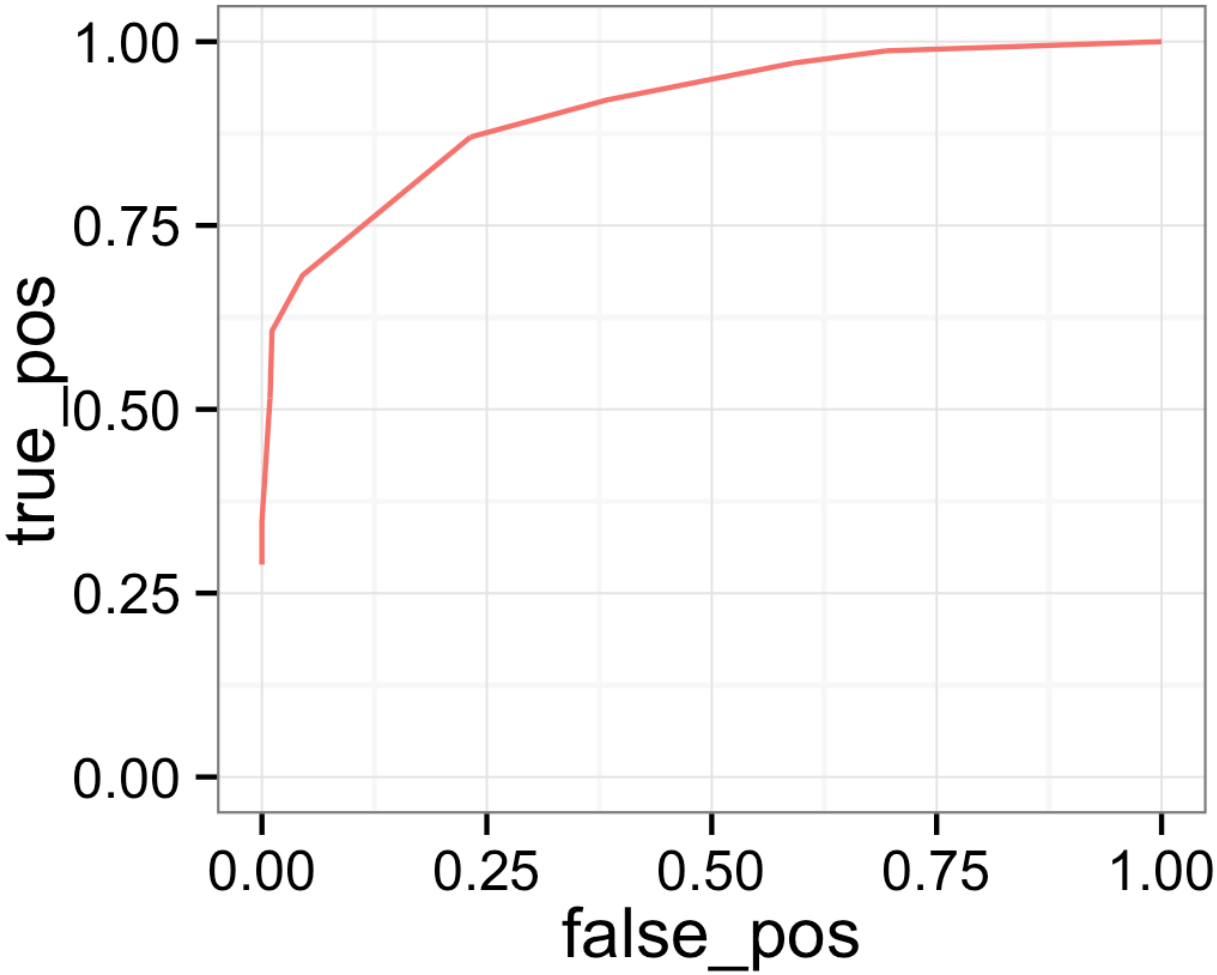
Sensitivity and specificity depend on a chosen cutoff



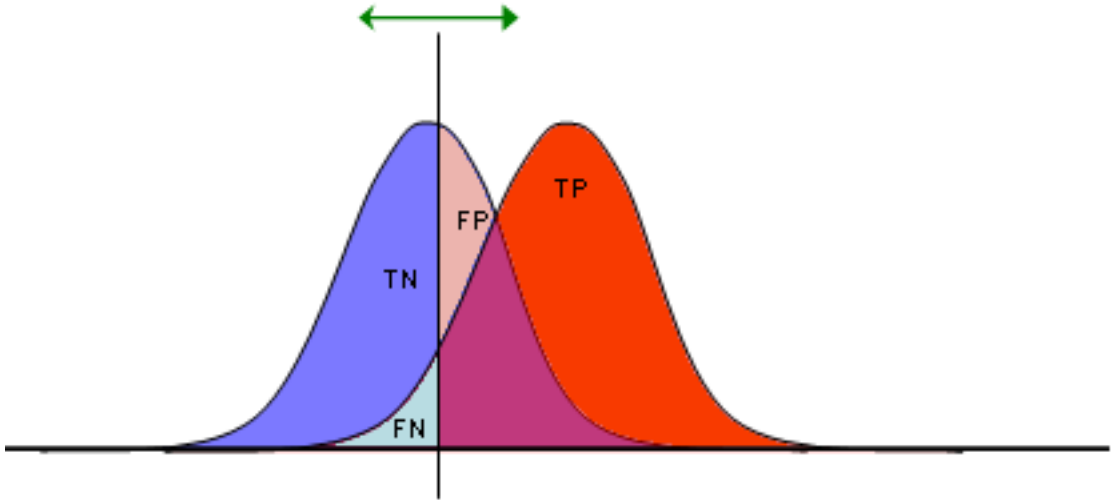
Sensitivity and specificity depend on a chosen cutoff



We usually plot the true pos. rate vs. the false pos. rate for all possible cutoffs



ROC curve
Receiver Operating Characteristic curve



TP	FP
FN	TN
1	1

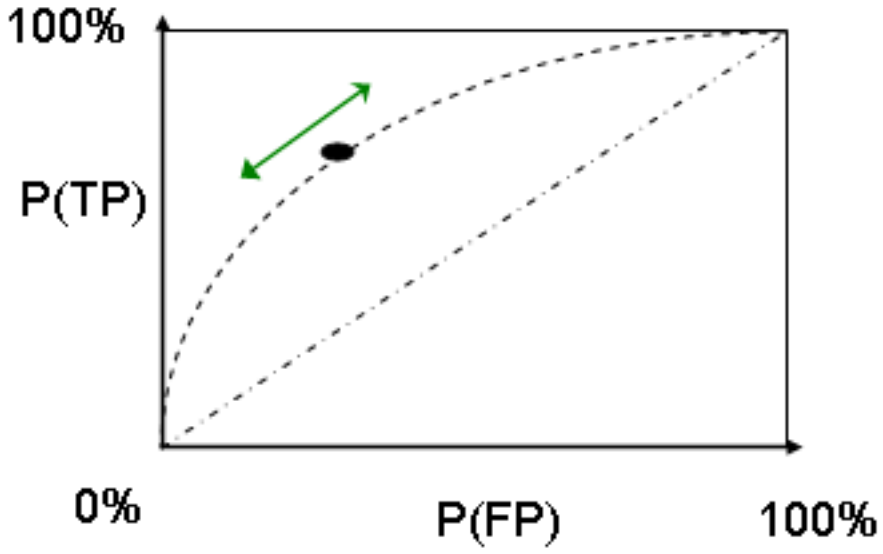
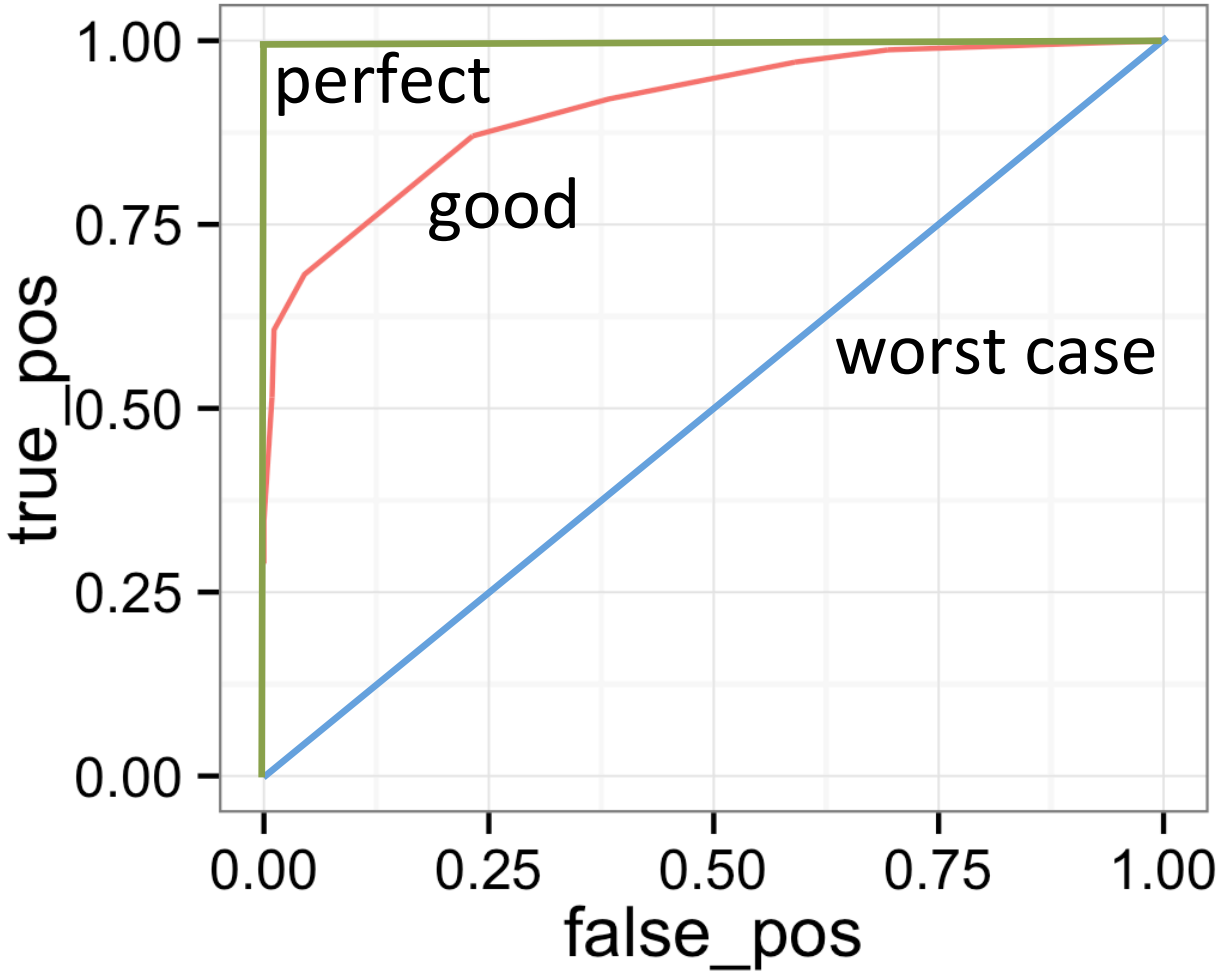
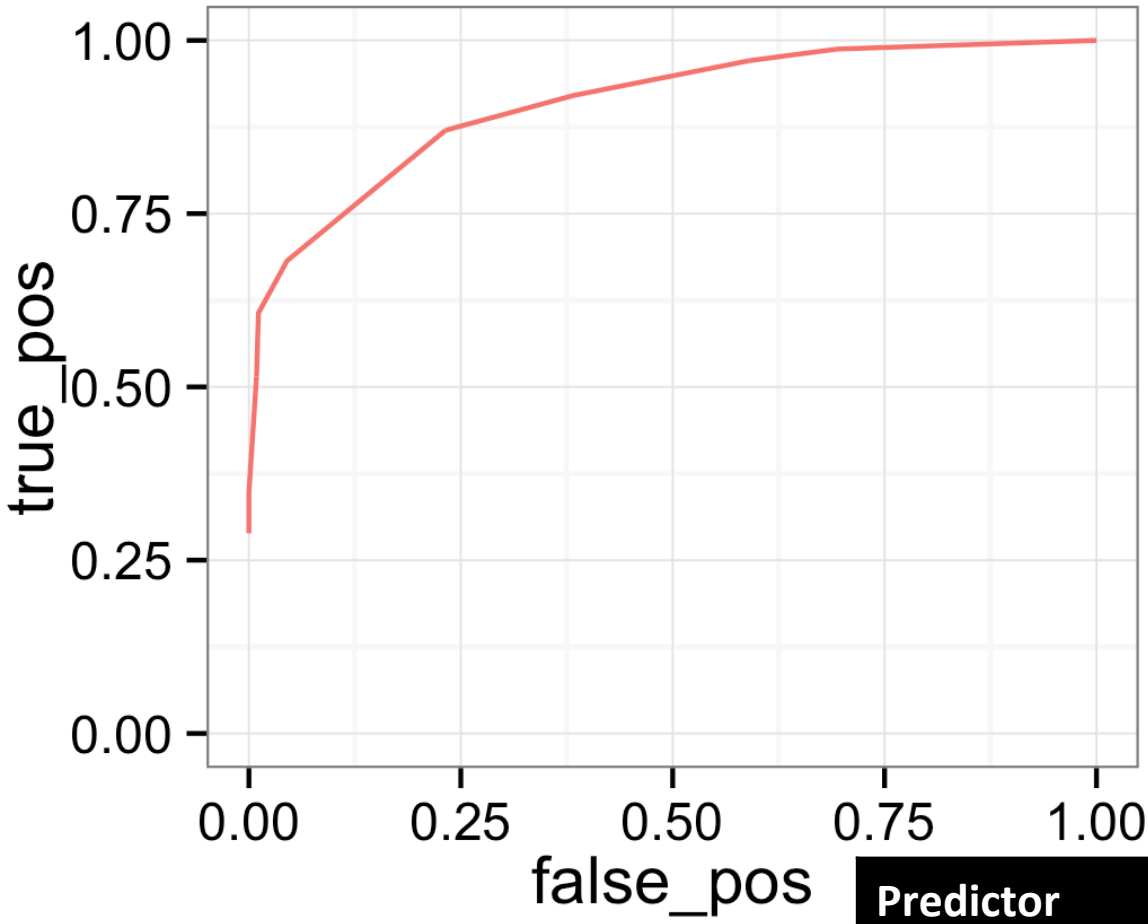


Image from: http://en.wikipedia.org/wiki/Receiver_operating_characteristic

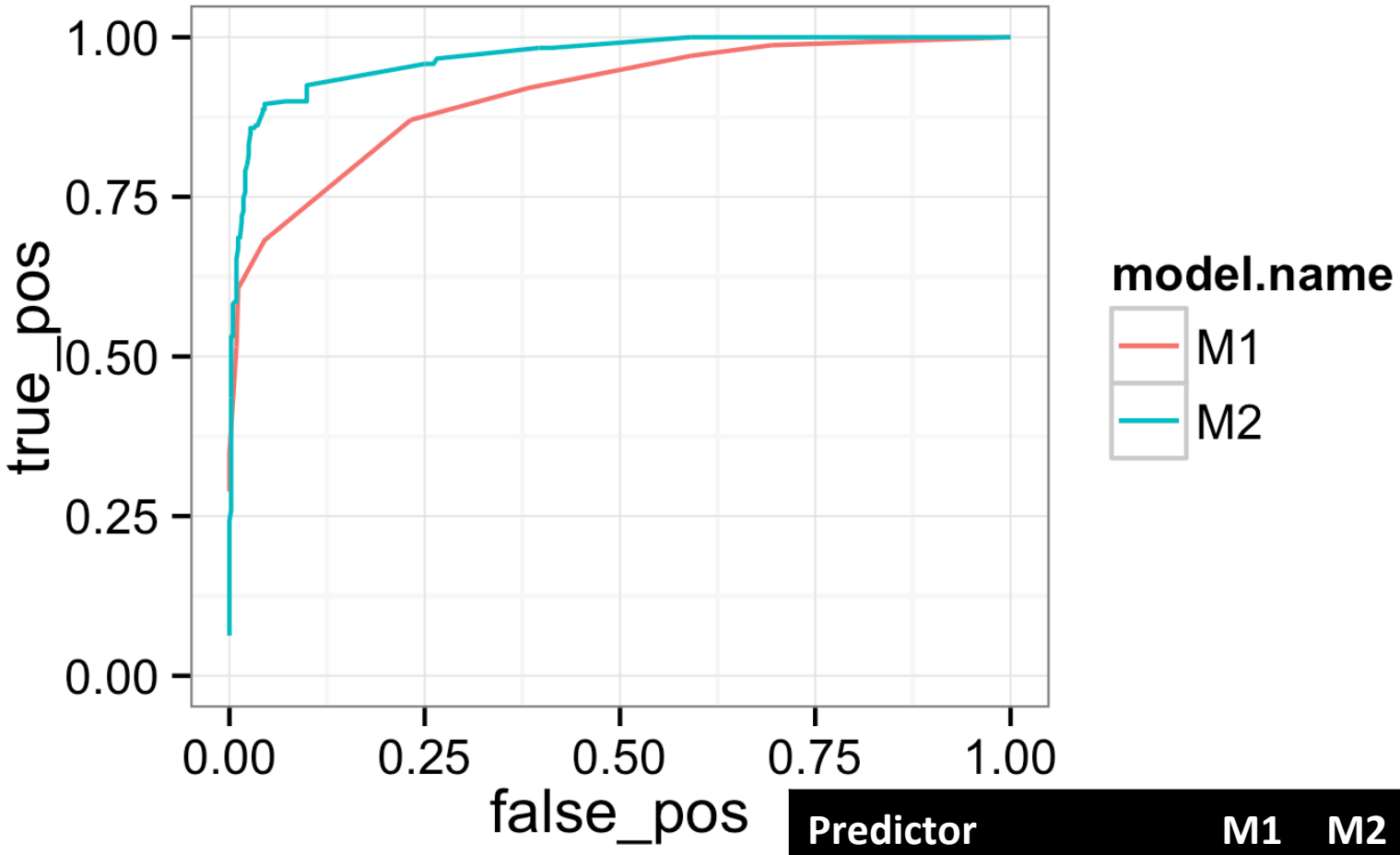
The area under the curve tells us how good a model's predictions are



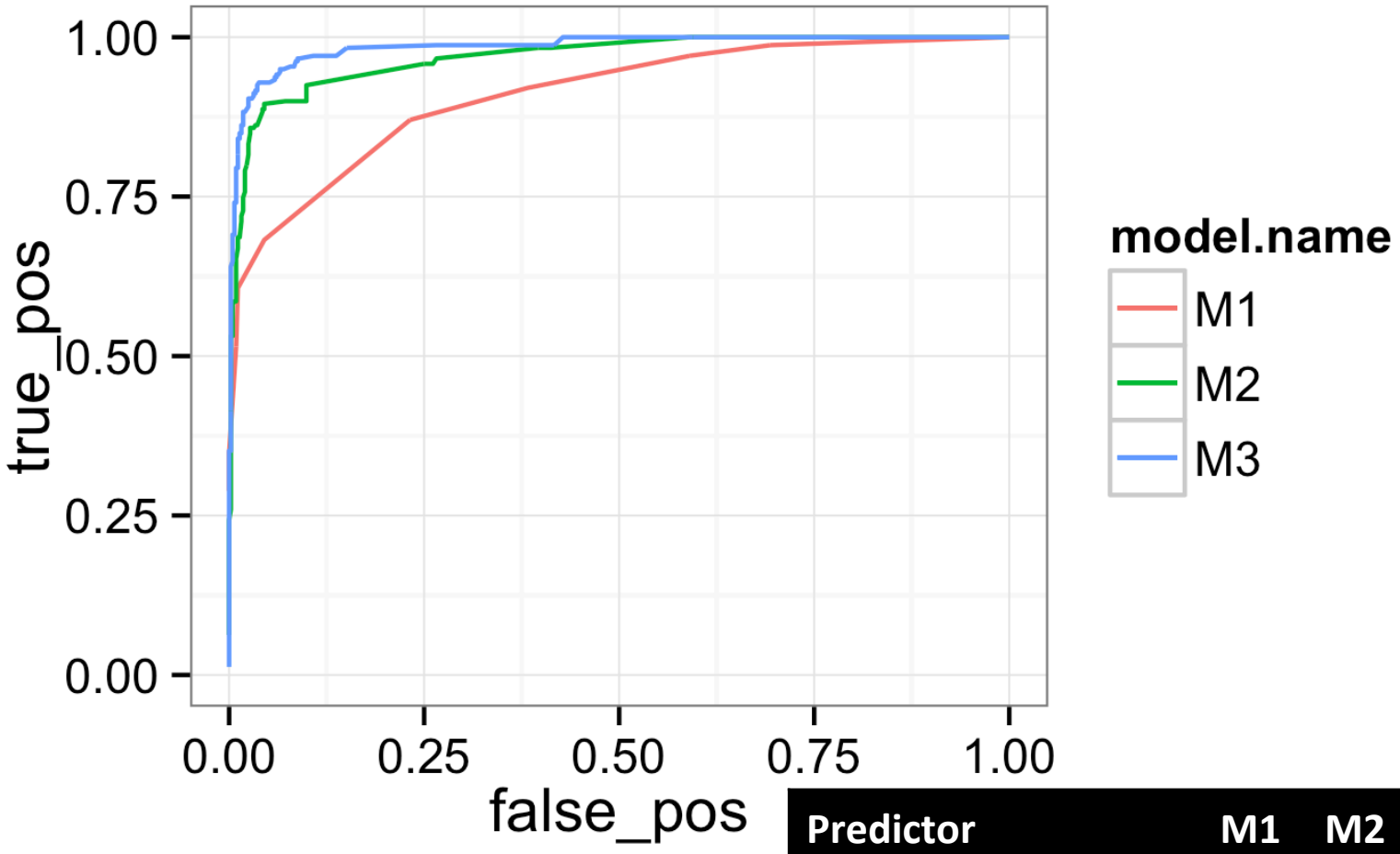
Let's look at the performance of several different models for the biopsy data set



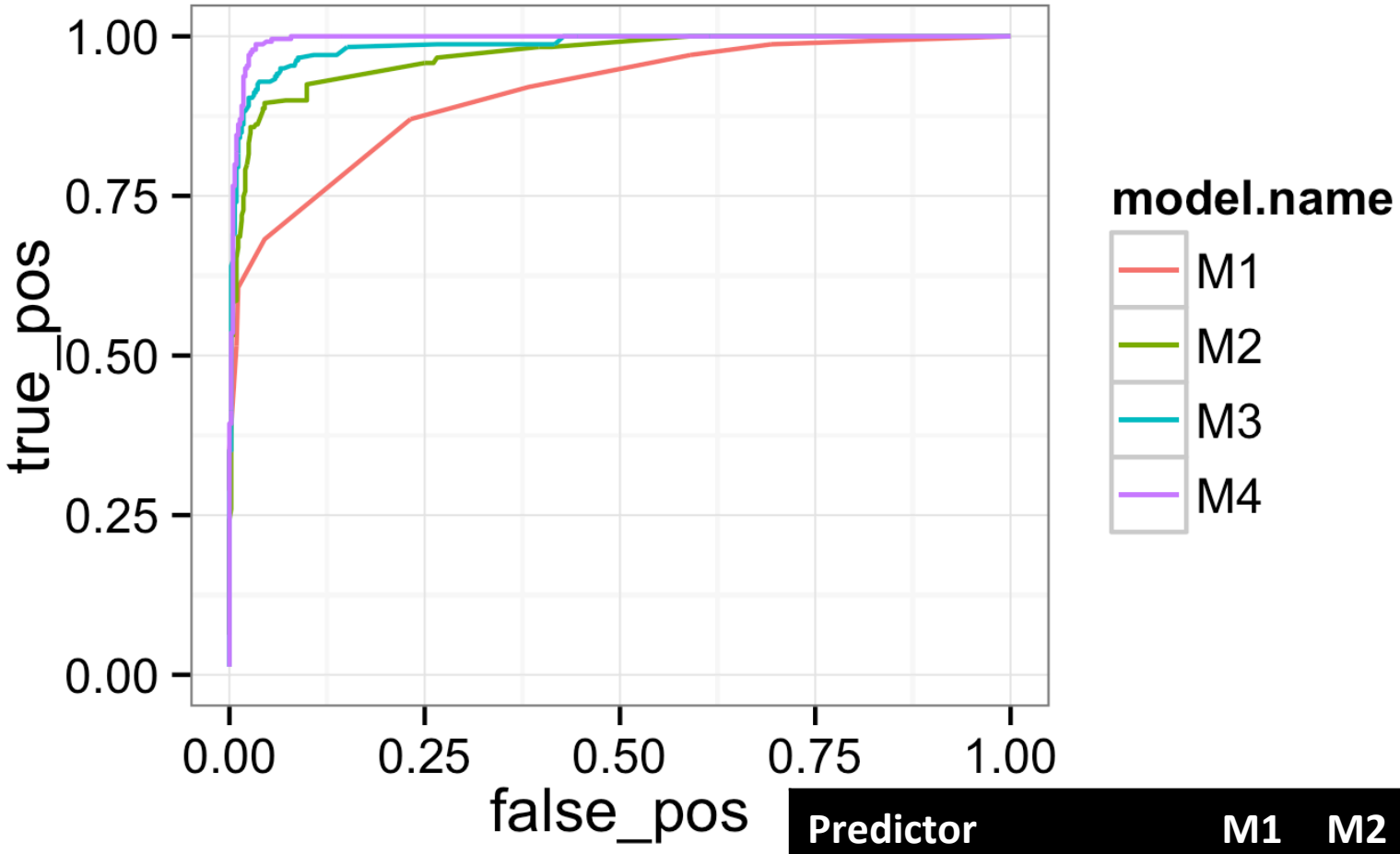
Predictor	M1
clump_thickness	✔
normal_nucleoli	
marg_adhesion	
bare_nuclei	
uniform_cell_shape	
bland_chromatin	



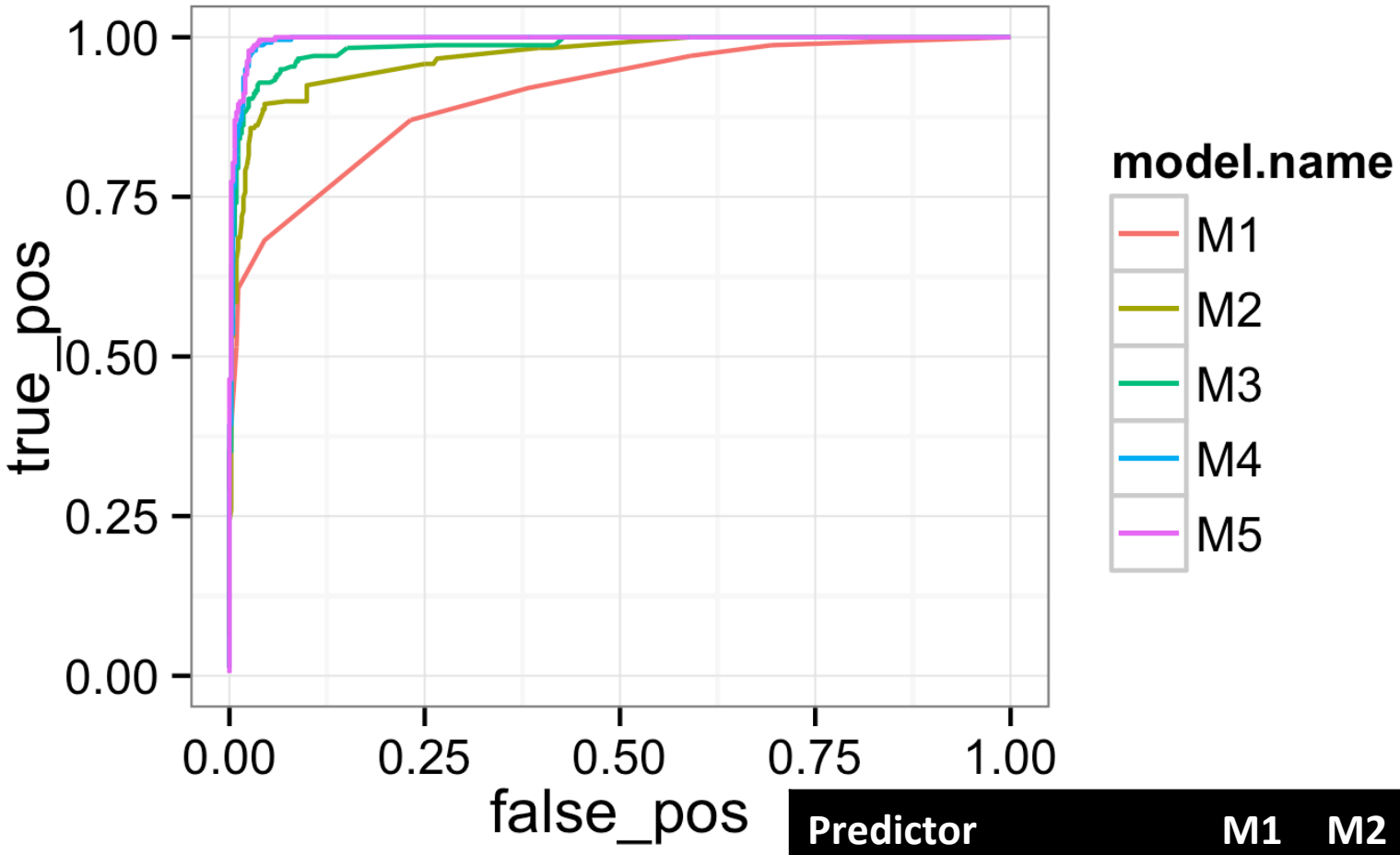
Predictor	M1	M2
clump_thickness	✔	✔
normal_nucleoli		✔
marg_adhesion		
bare_nuclei		
uniform_cell_shape		
bland_chromatin		



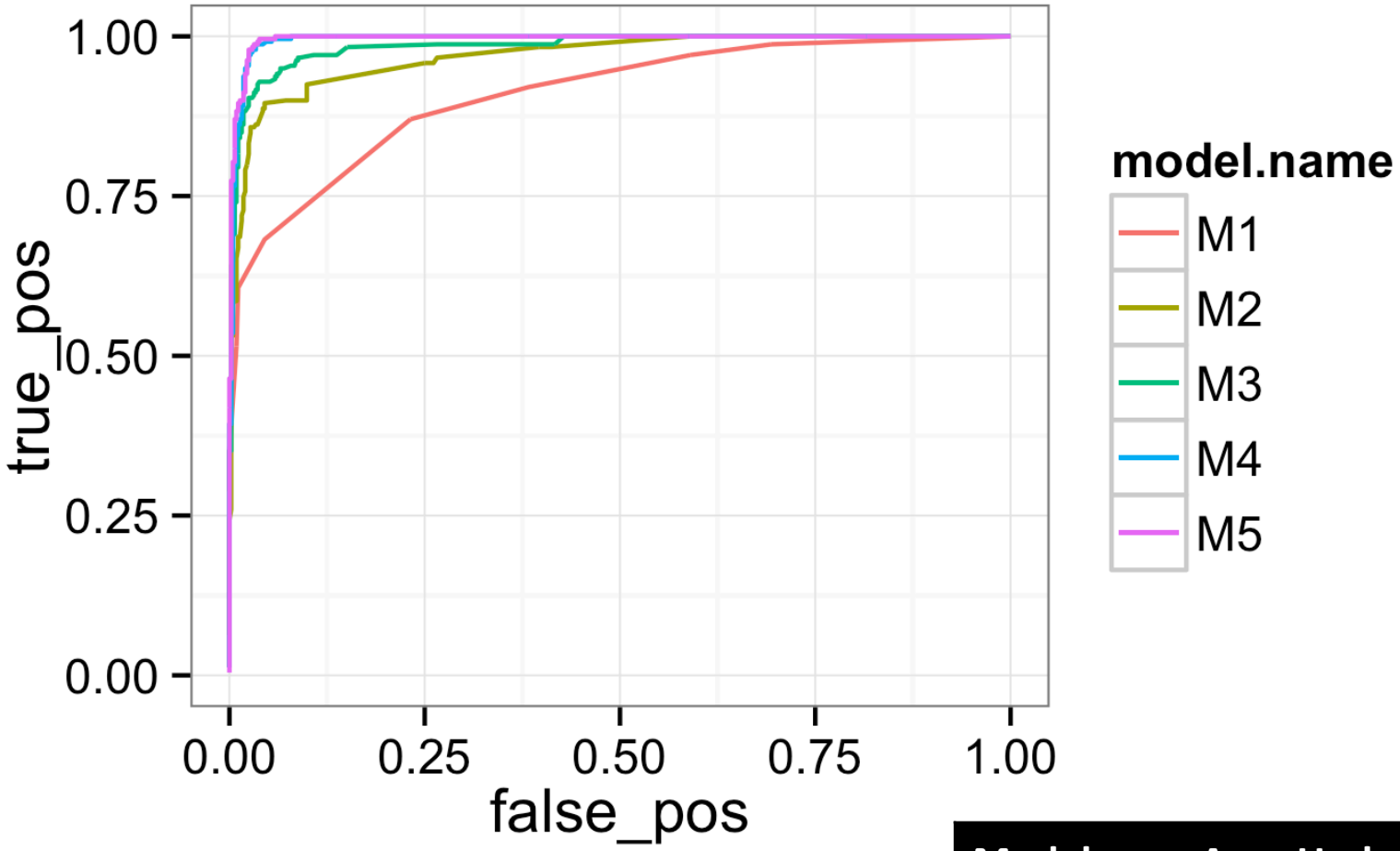
Predictor	M1	M2	M3
clump_thickness	✔	✔	✔
normal_nucleoli		✔	✔
marg_adhesion			✔
bare_nuclei			
uniform_cell_shape			
bland_chromatin			



Predictor	M1	M2	M3	M4
clump_thickness	✔	✔	✔	✔
normal_nucleoli		✔	✔	✔
marg_adhesion			✔	✔
bare_nuclei				✔
uniform_cell_shape				
bland_chromatin				

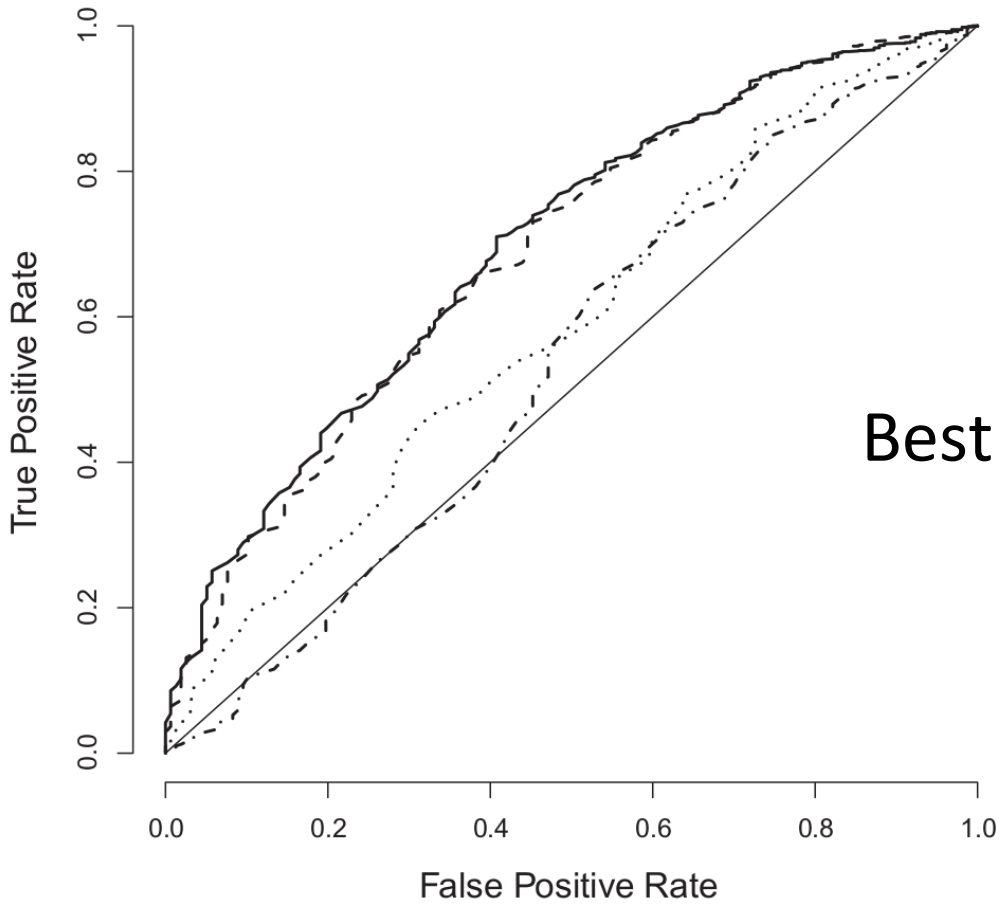


Predictor	M1	M2	M3	M4	M5
clump_thickness	✔	✔	✔	✔	✔
normal_nucleoli		✔	✔	✔	✔
marg_adhesion			✔	✔	✔
bare_nuclei				✔	✔
uniform_cell_shape					✔
bland_chromatin					✔



Model	Area Under Curve (AUC)
M1	0.940
M2	0.974
M3	0.985
M4	0.995
M5	0.996

Things usually look much worse in real life



Best AUC (solid line): 0.70

Keller, Mis, Jia, Wilke. Genome Biol. Evol. 4:80-88, 2012