Measures of Diagnostic Accuracy

Dr. Alison Smith

Lecturer, University of Leeds, UK.

A.F.C.smith@leeds.ac.uk

Learning Objectives

Measures of **diagnostic accuracy** tell you how well the test can classify individuals with or without a particular disease

Key diagnostic accuracy measures are derived from a **2 x 2 table**, which compares true disease status with test results

The below table presents sensitivity, specificity and predictive value results for low-dose CT:

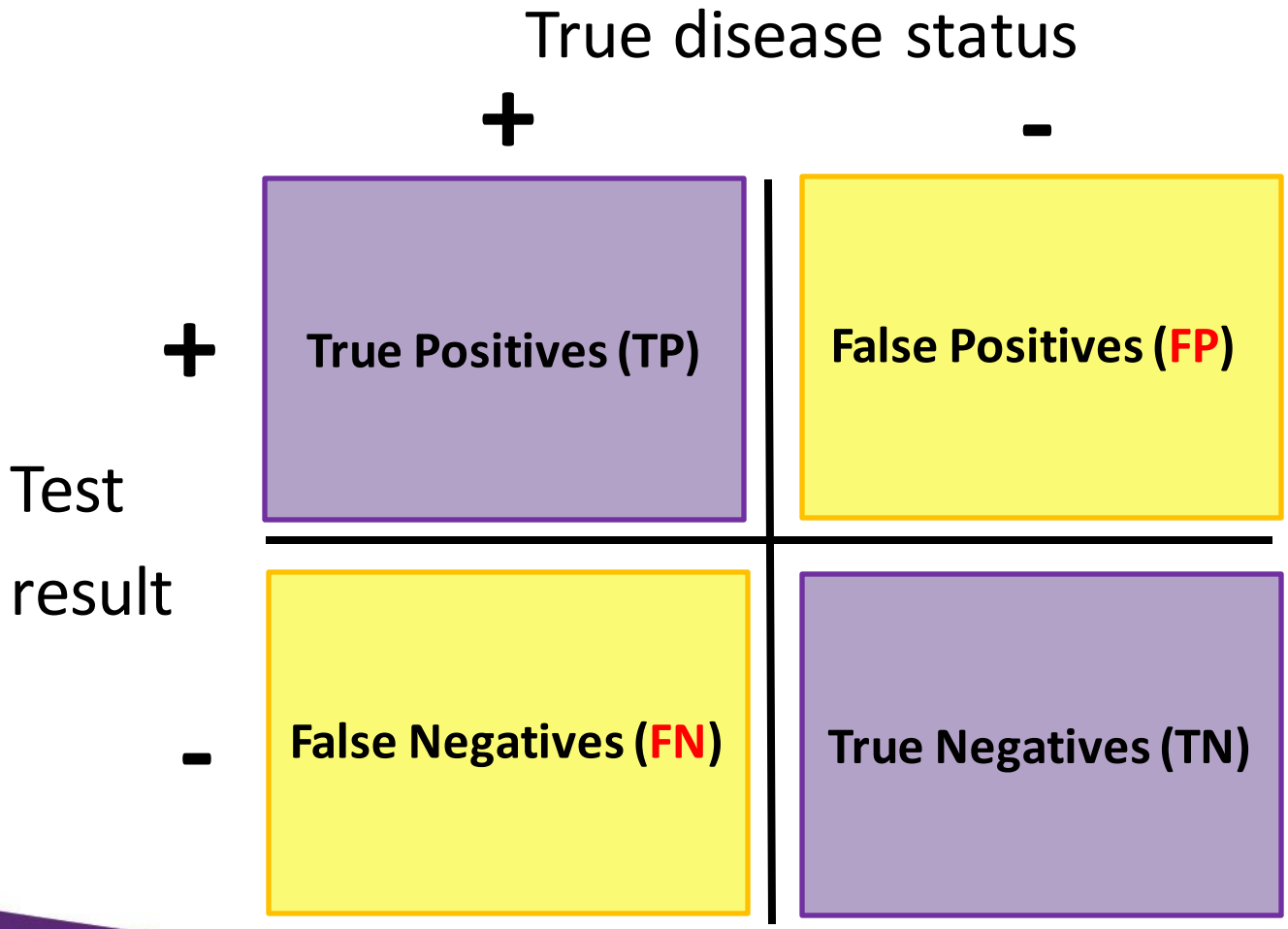
Diagnostic accuracy measure	Result
Sensitivity	93.8%
Specificity	73.4%
Positive predictive value	2.4 to 4.4%
Negative predictive value	99.9%

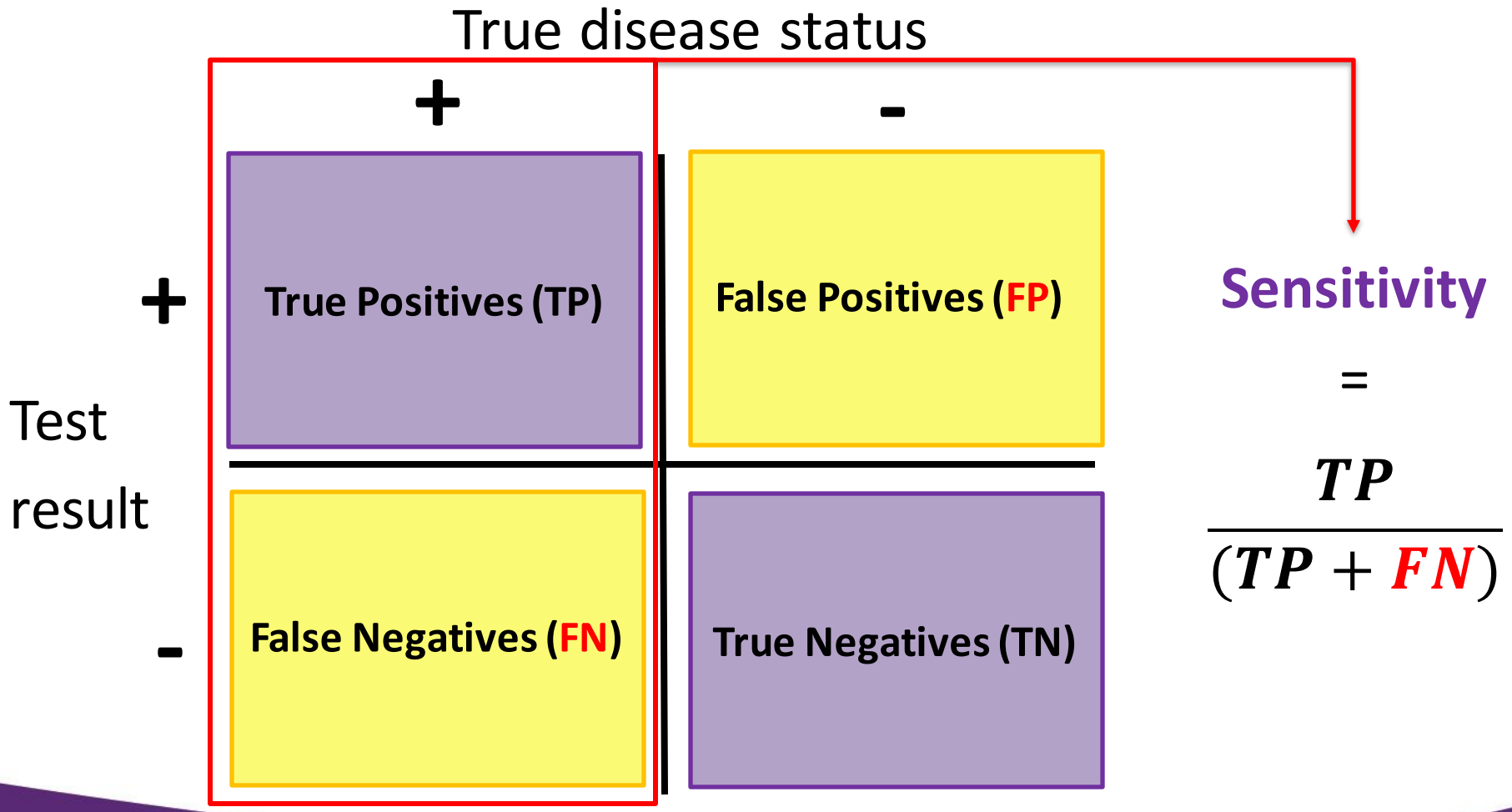
What do these numbers mean?

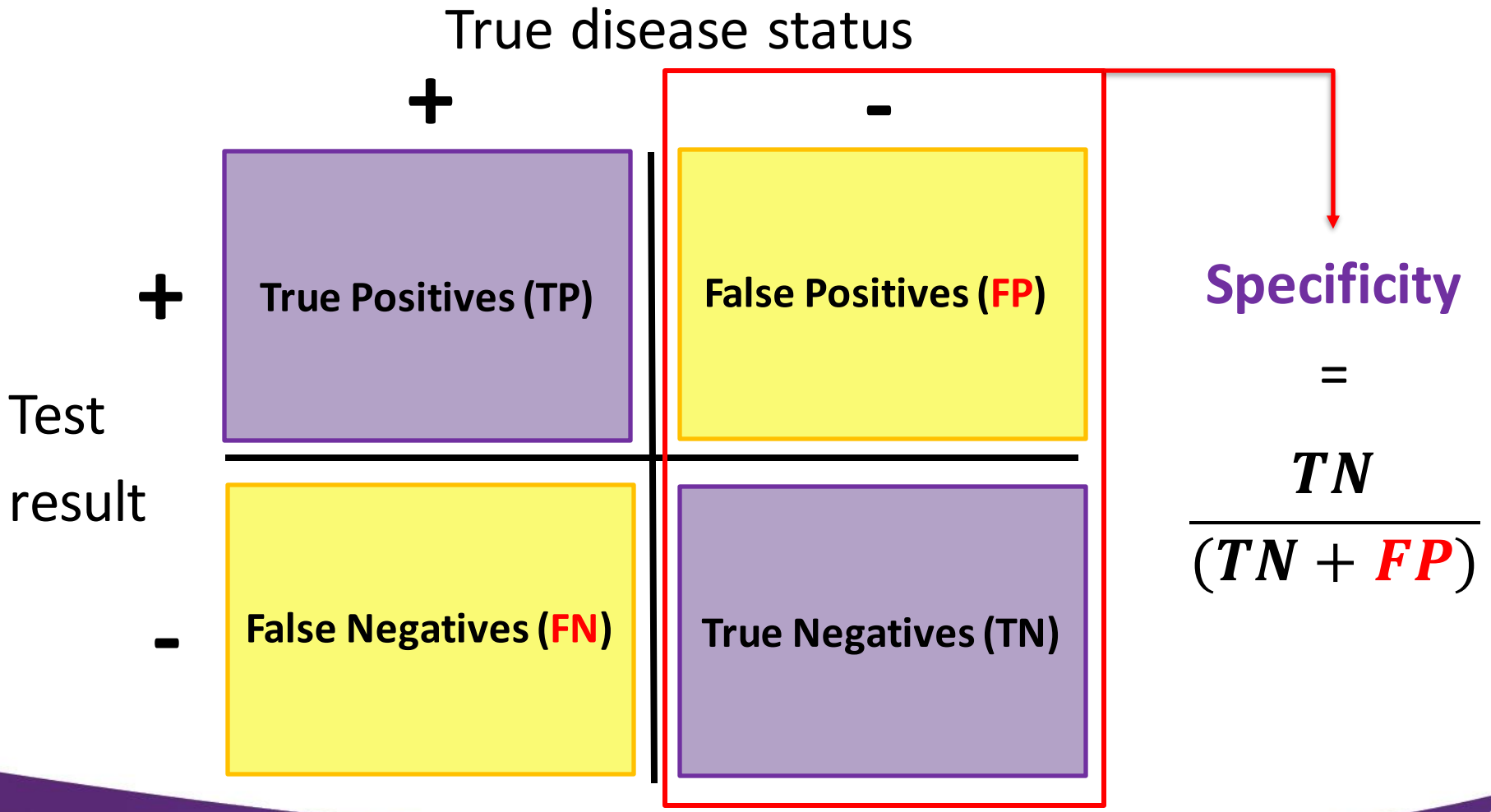
Accuracy values from: https://www.ncbi.nlm.nih.gov/books/NBK350016/pdf/Bookshelf_NBK350016.pdf

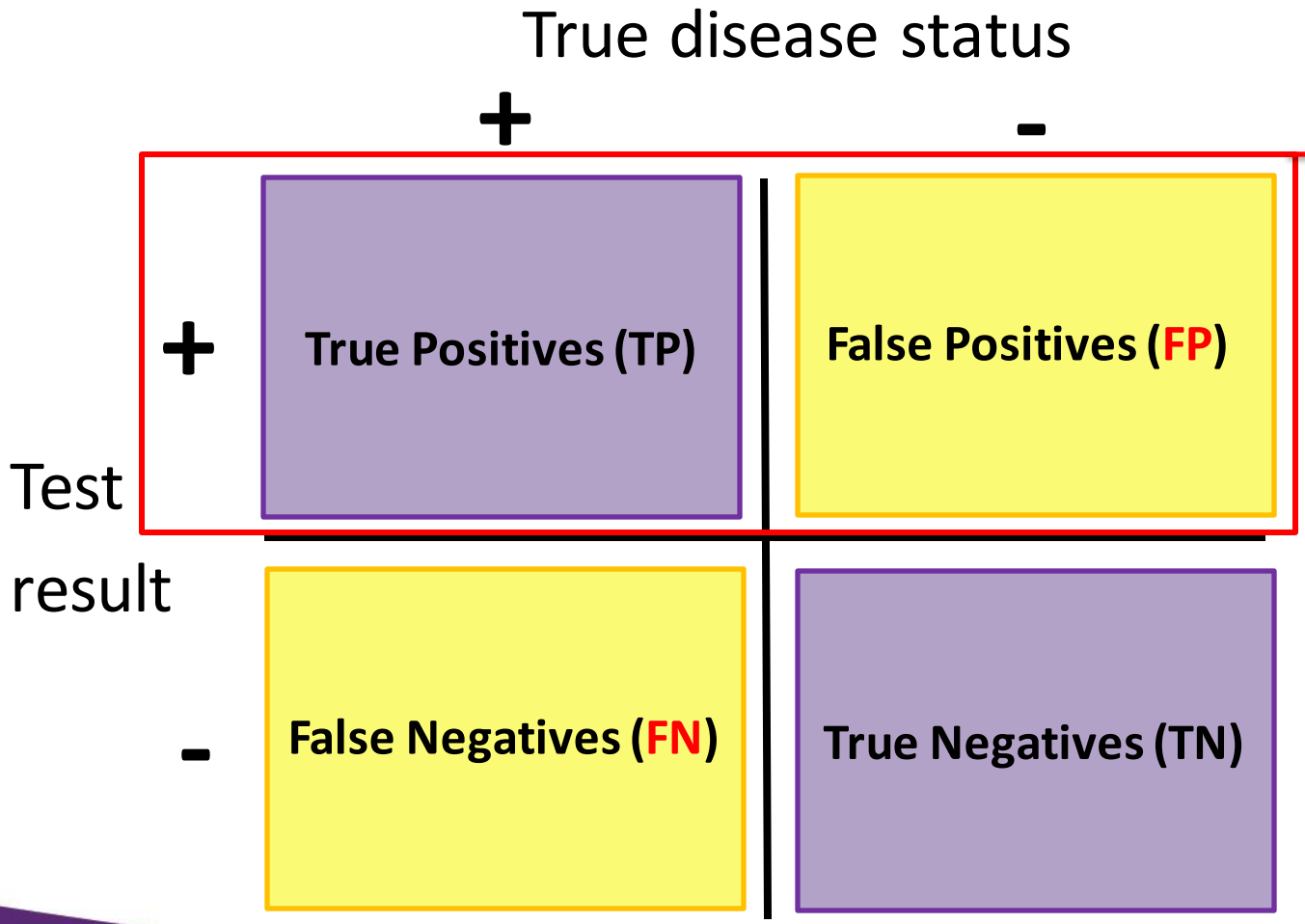
True Disease Status

- To evaluate the diagnostic accuracy of a test, we need to compare the test results to the 'truth'
- **Gold Standard**: the best test/method we have available for determining the disease status of an individual
- **Reference standard**: the best test we have available to estimate an individual's disease status


True disease status
+
-
Test result
+
-
True Positives (TP)
False Positives (FP)
False Negatives (FN)
True Negatives (TN)

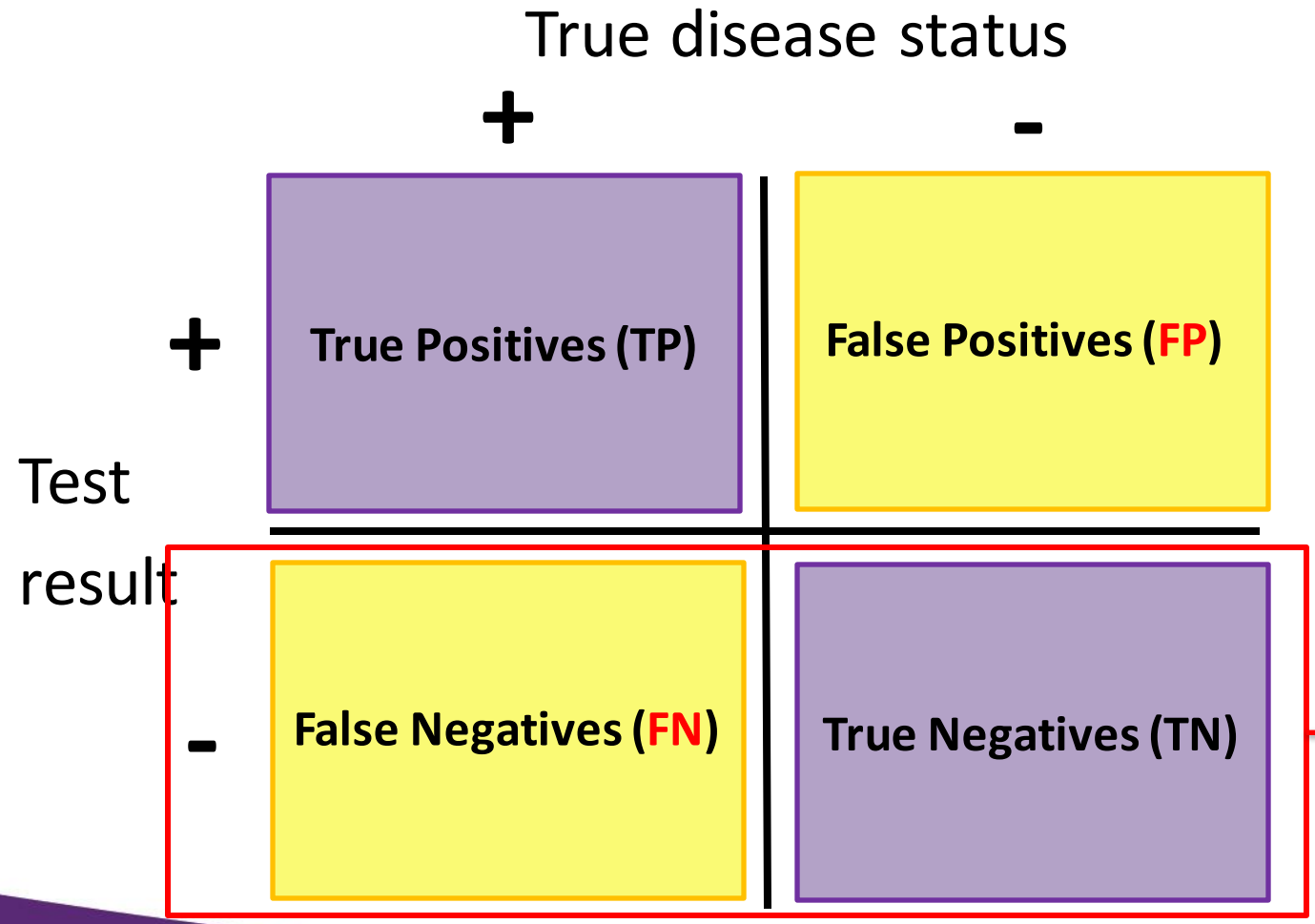

True disease status
+
-
Test result
+
-
True Positives (TP)
False Positives (FP)
False Negatives (FN)
True Negatives (TN)
Sensitivity
=
$\frac{TP}{(TP + FN)}$


True disease status
+
-
Test result
+
-
True Positives (TP)
False Positives (FP)
False Negatives (FN)
True Negatives (TN)
Specificity
=
$\frac{TN}{(TN + FP)}$



Positive predictive value

$$= \frac{TP}{(TP + FP)}$$



Negative predictive value

$$= \frac{TN}{(TN + FN)}$$

The below table presents sensitivity, specificity and predictive value results for low-dose CT:

Diagnostic accuracy measure	Result
Sensitivity	93.8%
Specificity	73.4%
Positive predictive value	2.4 to 4.4%
Negative predictive value	99.9%

What do these numbers mean?

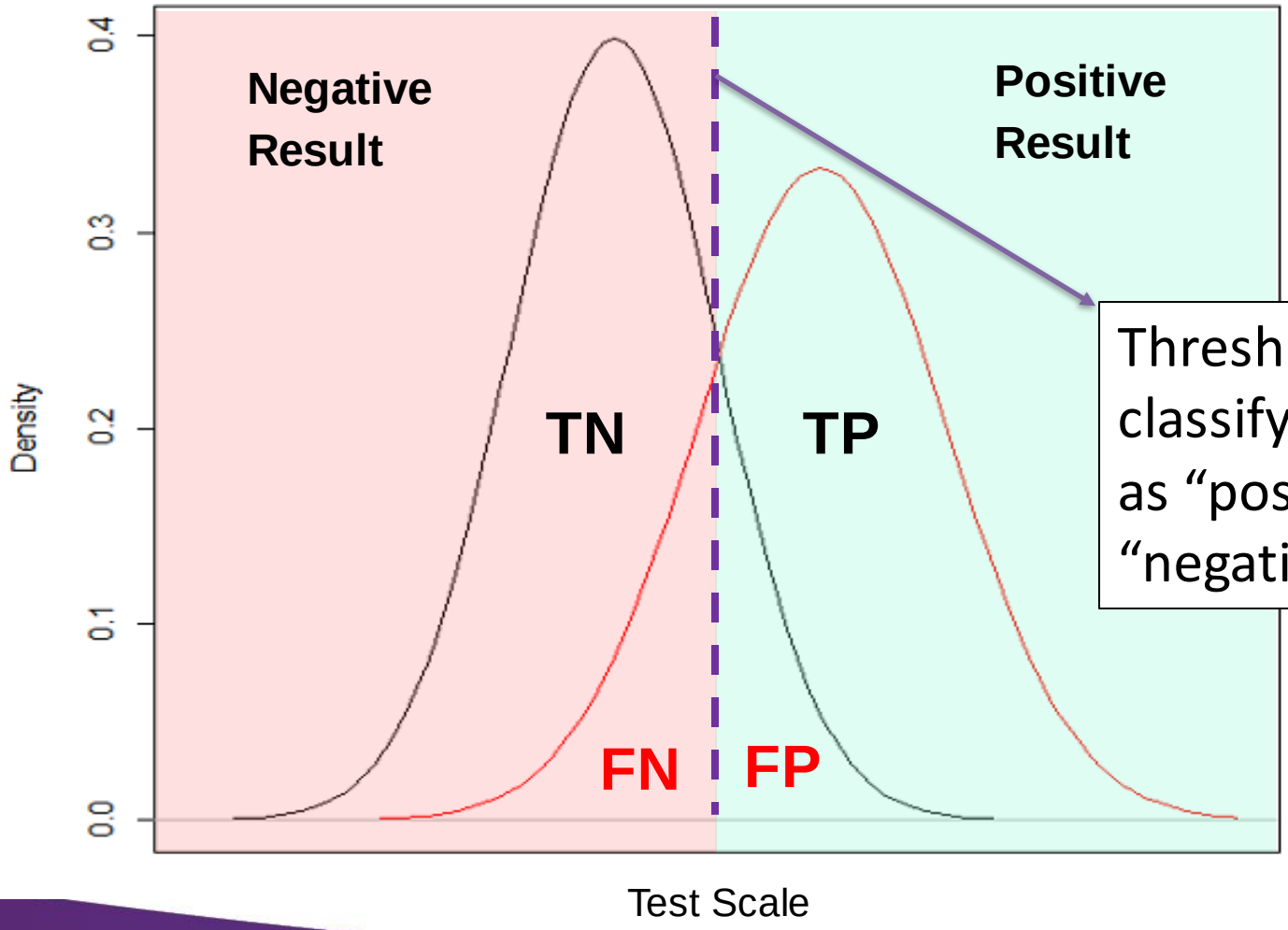
Accuracy values from: https://www.ncbi.nlm.nih.gov/books/NBK350016/pdf/Bookshelf_NBK350016.pdf

Test diagnostic threshold

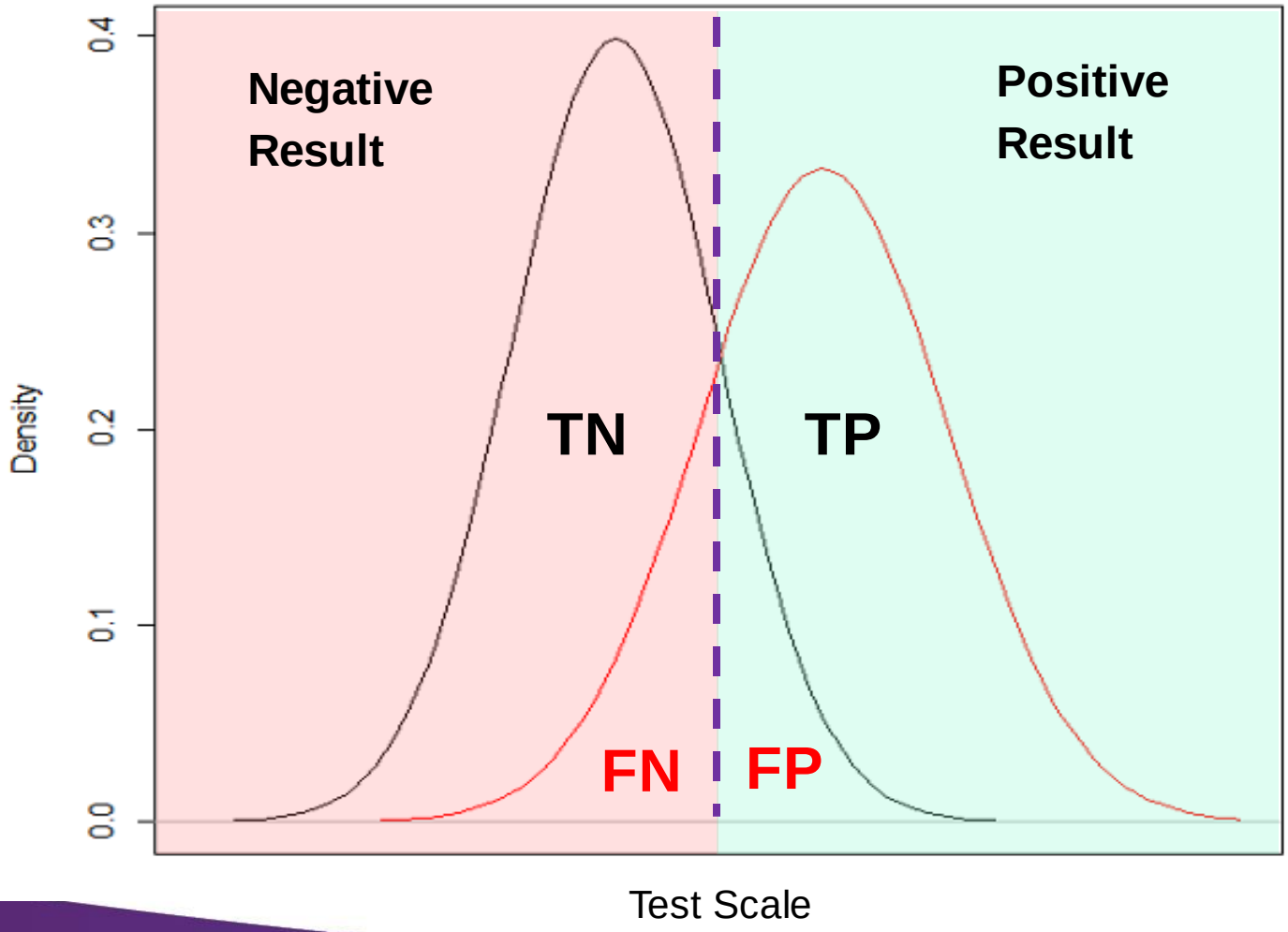
Test results are classified as ‘positive’ or ‘negative’ against a specified **threshold value** or positivity criterion.

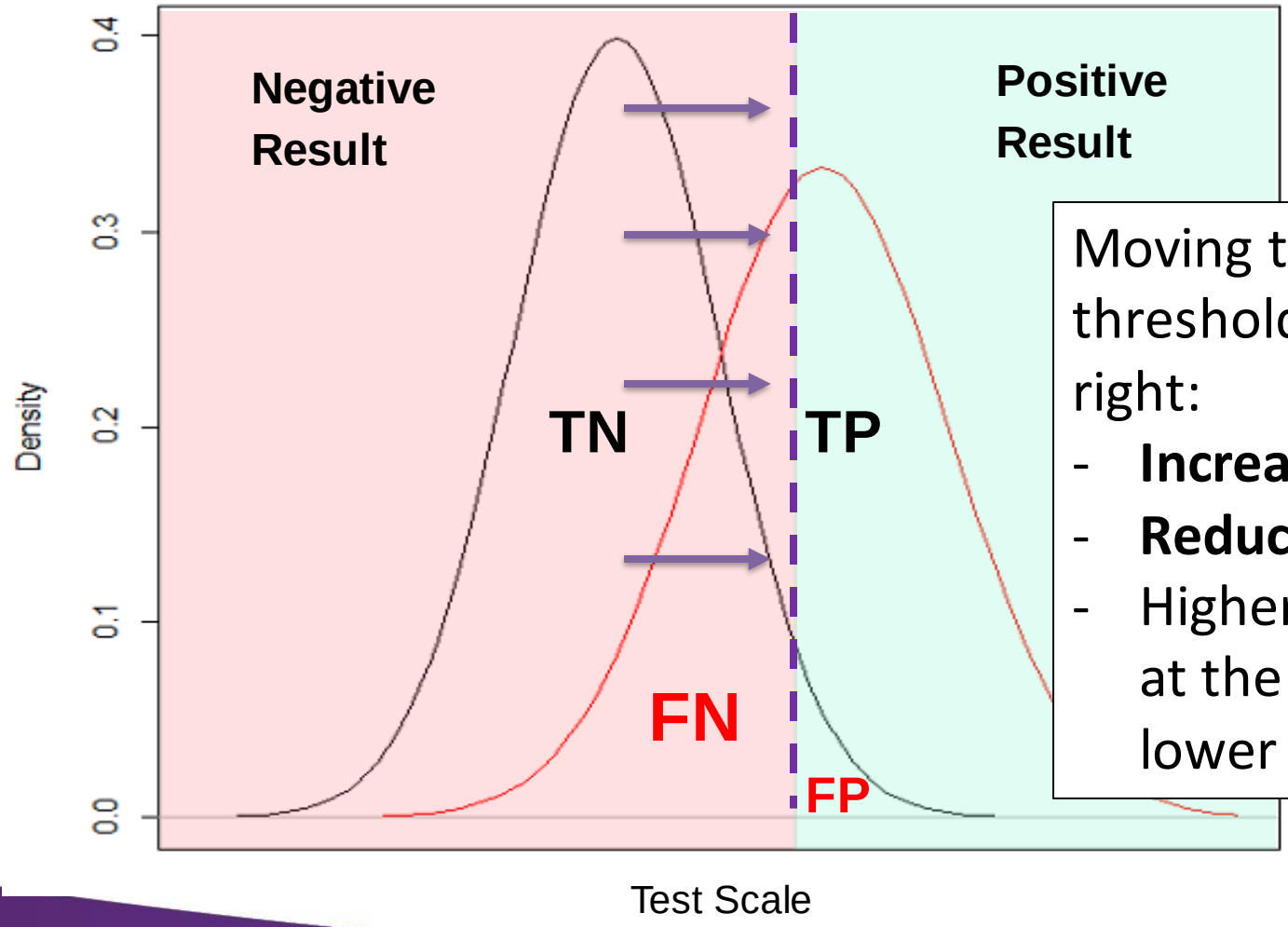
E.g. for tests based on the quantitative measurement of a biomarker, the threshold will be a given numerical value.

If the threshold changes, the diagnostic accuracy will also change.



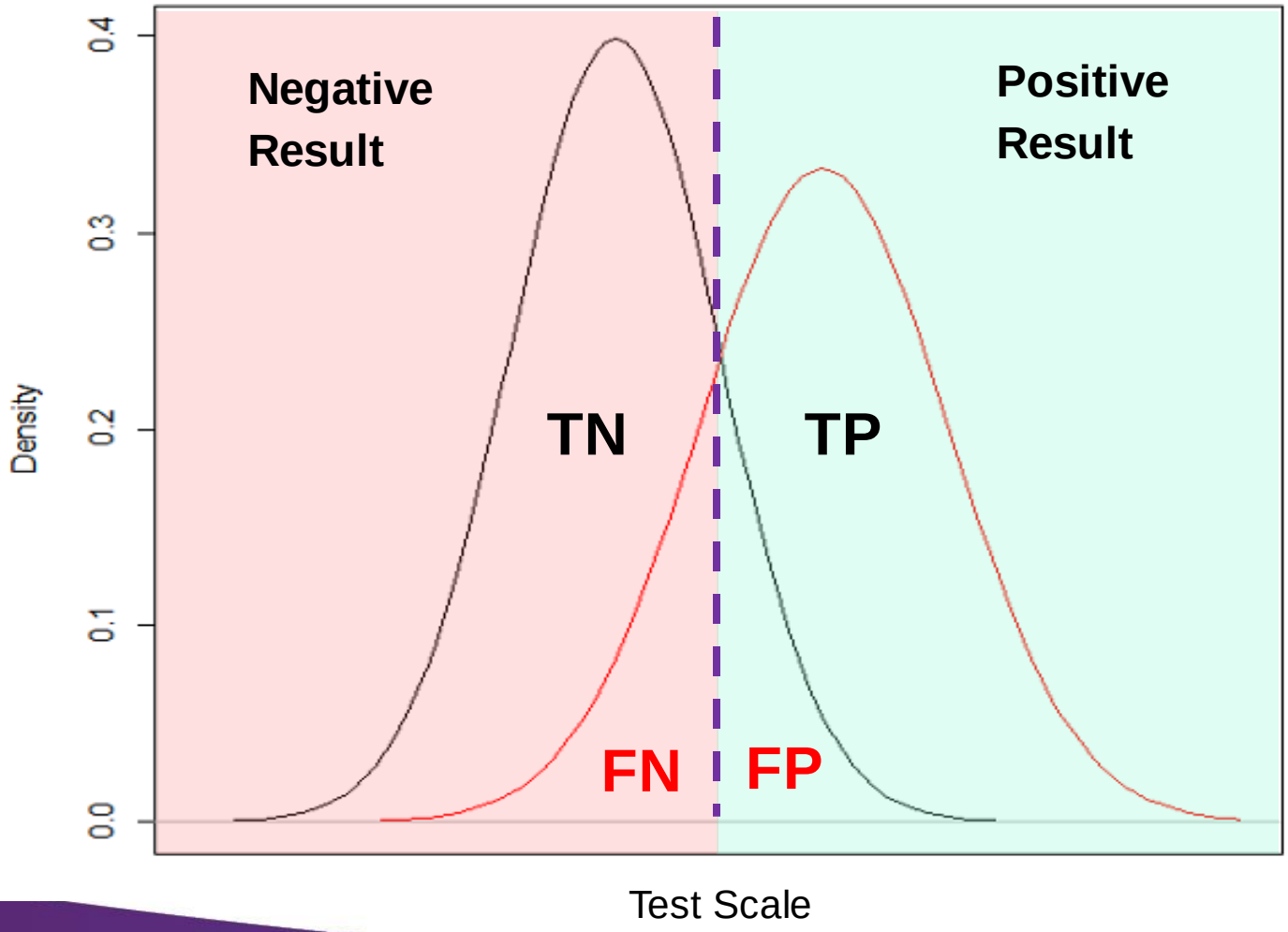
Threshold used to classify test results as “positive” vs. “negative”

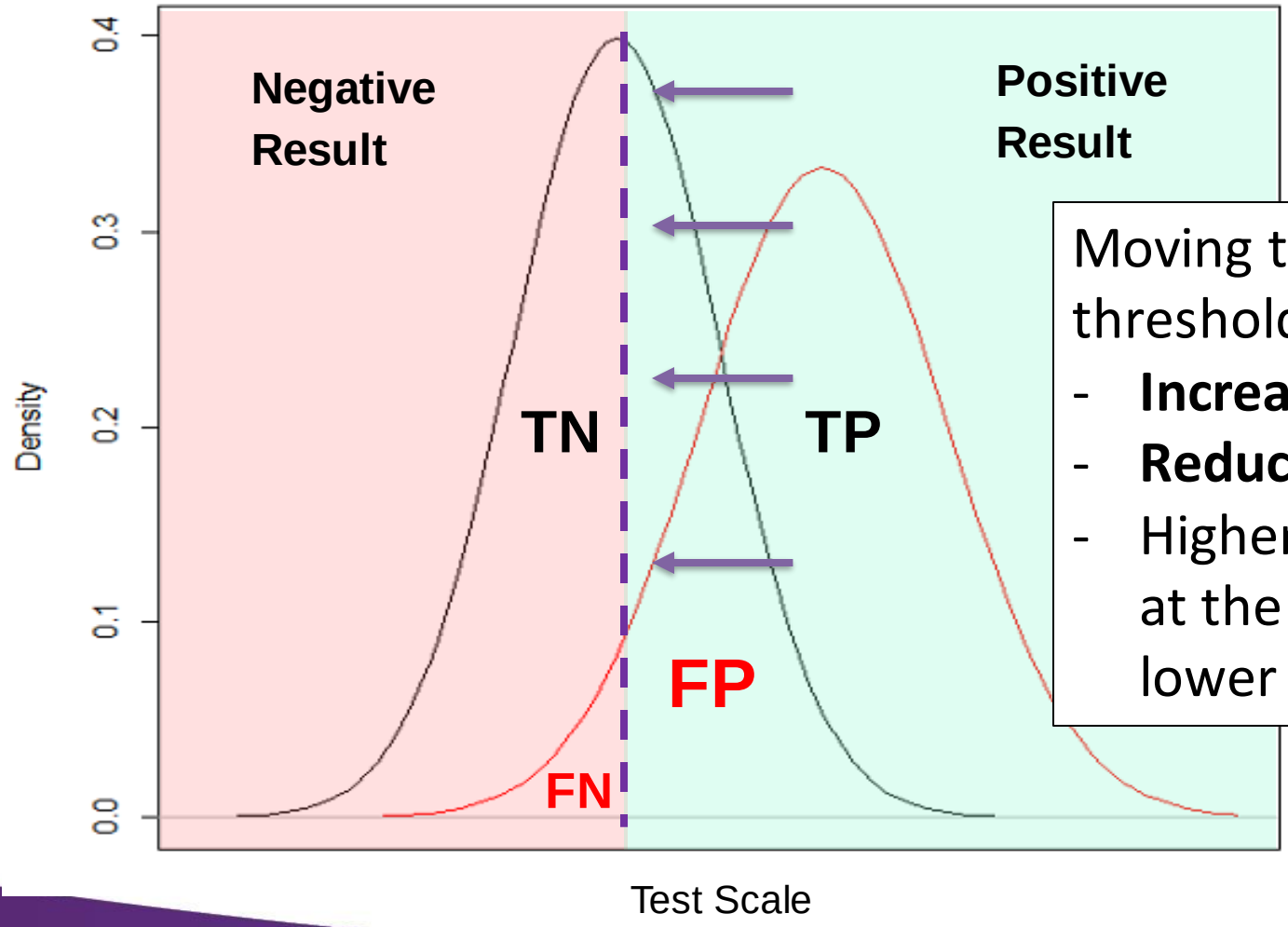

Negative
Result
Positive
Result
TN
TP
FN
FP
Density
0.0
0.1
0.2
0.3
0.4
Test Scale



Moving the threshold to the right:

- **Increases FN**
- **Reduces FP**
- Higher specificity, at the cost of lower sensitivity


Negative
Result
Positive
Result
TN
TP
FN
FP
Density
0.0
0.1
0.2
0.3
0.4
Test Scale



Moving the threshold to the left:

- **Increases FP**
- **Reduces FN**
- Higher sensitivity, at the cost of lower specificity

Prevalence and spectrum of disease

Prevalence is defined as the proportion of a population who have a particular disease or condition. The **spectrum of disease** describes the proportion of patients with mild vs. severe illness.

Changing the **prevalence** of disease can significantly affect predictive values (**PPV** and **NPV)** but does not affect sensitivity or specificity.

HOWEVER, changes in the **spectrum of disease** (which often occur together with changes in the prevalence) can significantly affect **sensitivity** and/or **specificity**.

Example

In the Excel file called ‘test accuracy’, the accuracy of a hypothetical test is compared across two settings: ‘A&E’ and ‘GP’ (each shown in separate worksheets).

For each setting, we have given you the estimated disease prevalence in this patient population, and the test sensitivity and specificity.

Have a look in your own time…

Imagine you carry out the test on 100 people in each setting. Work through the flowcharts and fill in the missing numbers to **calculate the PPVs and NPVs.**

What is the PPV and NPV of the test in each clinical setting? Do you think this is a useful test to use in the GP or A&E settings?

Summary

- Diagnostic accuracy values can be calculated from the 2x2 table
- Sensitivity and specificity consider the accuracy of a test from the perspective of patients with the disease (sensitivity) and without the disease (specificity)
- PPV and NPV consider the accuracy of a test from the perspective of patients with a positive test result (PPV) or negative test result (NPV)
- Key factors to consider include: the diagnostic threshold, disease prevalence, and the spectrum of disease