

Diagnostic ability assessment of quantitative medical tests

Sara Perez-Jaume

February 2026



UNIVERSITAT DE
BARCELONA

Contents

- 1 Introduction
- 2 Continuous tests
- 3 ROC curves
- 4 Optimum threshold estimation
- 5 Real-case example
- 6 Conclusions

Contents

- 1 Introduction
- 2 Continuous tests
- 3 ROC curves
- 4 Optimum threshold estimation
- 5 Real-case example
- 6 Conclusions

Introduction

- Accurate diagnosis of a health condition is of utmost importance.
- Appropriate treatment depends on knowing the condition affecting the patient.
- Research studies examine the ability of diagnostic tests to discriminate between individuals with or without the condition of interest.

Introduction

- A diagnostic test is a measurement taken from an individual with the aim of diagnosing a disease or condition.
- Binary and continuous diagnostic tests.

Binary tests

- Some diagnostic tests are binary (dichotomous).
 - Antigen test for COVID-19 (+ versus -)
- Binary tests are sometimes the result of dichotomizing a continuous test by means of a threshold.
 - Cholesterol levels in blood ($\geq 200\text{mg/dL}$, which indicates cardiovascular risk, versus $< 200\text{mg/dL}$)

Binary tests: Notation

- Disease status (true condition, gold standard):
 - $D \rightarrow$ Diseased (condition present)
 - $\bar{D} \rightarrow$ Non-diseased (condition absent, healthy)
- Diagnostic test result:
 - $Y \rightarrow$ Positive, + (test results indicate the condition is present)
 - $\bar{Y} \rightarrow$ Negative, – (test results indicate the condition is absent)

Decision matrix

Test	True condition		Total
	D	$\bar{D}$	
Y	TP	FP	TP + FP
$\bar{Y}$	FN	TN	FN + TN
Total	TP + FN	FP + TN	n

Table 1: 2×2 general decision matrix of a binary test

TP: true positives, TN: true negatives,
FP: false positives, FN: false negatives.

[Pepe, 2003, Zhou et al., 2011, Nakas et al., 2023]

Sensitivity and specificity

- **Sensitivity:** the probability of detecting the condition when it is present:

$$\text{Sens} = P(Y|D) = \frac{TP}{TP + FN}.$$

- **Specificity:** the probability of excluding the condition when it is absent:

$$\text{Spec} = P(\bar{Y}|\bar{D}) = \frac{TN}{FP + TN}.$$

Contents

- 1 Introduction
- 2 Continuous tests
- 3 ROC curves
- 4 Optimum threshold estimation
- 5 Real-case example
- 6 Conclusions

Continuous tests

- In many situations, the diagnostic test is continuous rather than binary.
- Example: Prostate-Specific Antigen (PSA) (ng/mL) is used to distinguish between individuals with or without prostate cancer.

Continuous tests

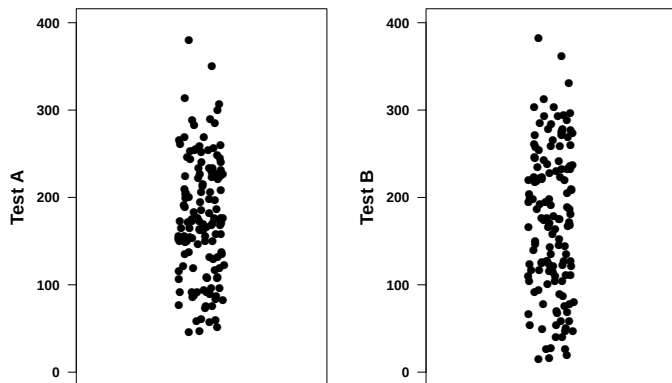


Figure 1: Two examples: quantitative diagnostic test

Continuous tests

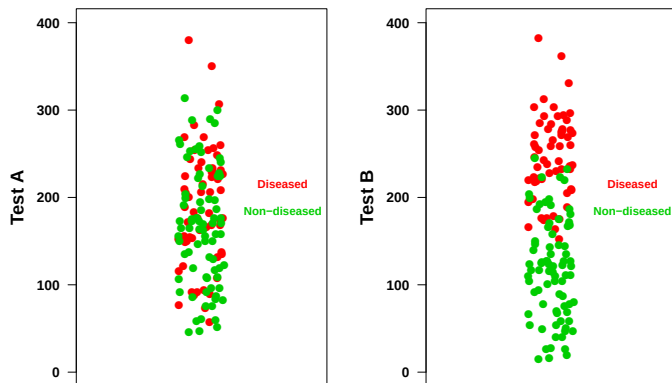


Figure 1: Two examples: quantitative diagnostic test

Continuous tests: Notation

- Disease status (true condition, gold standard):
 - $D \rightarrow$ Diseased (condition present)
 - $\bar{D} \rightarrow$ Non-diseased (condition absent, healthy)
- Continuous diagnostic test: X .
- We assume that higher values of X are indicative of disease (if not, change the labels!).

Decision matrix

- Threshold for X : c .
- We consider: $X \geq c \rightarrow$ positive, $X < c \rightarrow$ negative.

Test	True condition		Total
	D	$\bar{D}$	
$X \geq c$	TP(c)	FP(c)	TP(c) + FP(c)
$X < c$	FN(c)	TN(c)	FN(c) + TN(c)
Total	TP(c) + FN(c)	FP(c) + TN(c)	n

Table 2: 2×2 general decision matrix of a continuous test

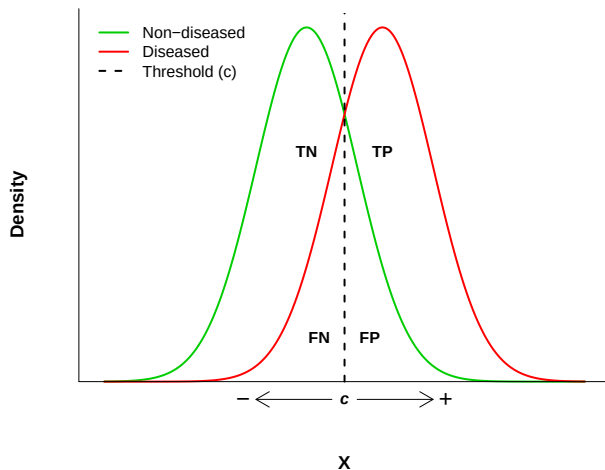


Figure 2: Example of a continuous diagnostic test

Sensitivity and specificity

Then, the sensitivity and specificity of X depend on c :

$$\text{Sens}(c) = P(X \geq c \mid D),$$

$$\text{Spec}(c) = P(X < c \mid \bar{D}).$$

Contents

- 1 Introduction
- 2 Continuous tests
- 3 ROC curves**
- 4 Optimum threshold estimation
- 5 Real-case example
- 6 Conclusions

ROC curves

- The ROC¹ curve is defined as the whole set of possible sensitivities and specificities we can obtain from the dichotomization of X with different thresholds [Pepe, 2003], that is,

$$\{(1 - \text{Spec}(c), \text{Sens}(c)), c \in (-\infty, \infty)\}$$

- Plot: x-axis $\rightarrow 1 - \text{Spec}$; y-axis $\rightarrow \text{Sens}$.
- The ROC curve is a monotone increasing function.
- It summarizes the ability of the test to discriminate among the two true states.

¹Receiver Operating Characteristic

ROC curves

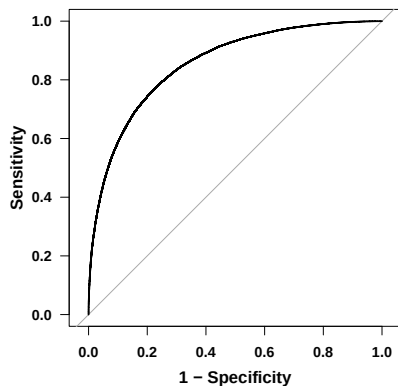


Figure 3: Example of an ROC curve

Area under the ROC curve (AUC)

- It is often useful to summarize the ROC curve by a single number [Pepe, 2003, Zhou et al., 2011].
- Area under the ROC curve $\rightarrow$ AUC.
- $0.5 < \text{AUC} \leq 1$.

AUC

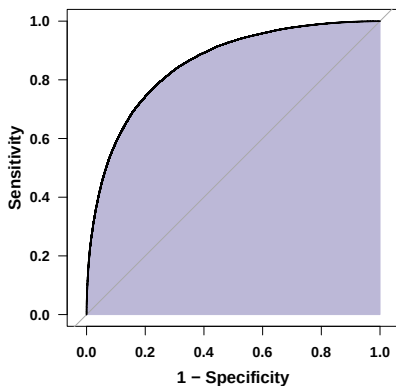


Figure 4: Example of an ROC curve with AUC in grey

$$\text{AUC} = 1$$

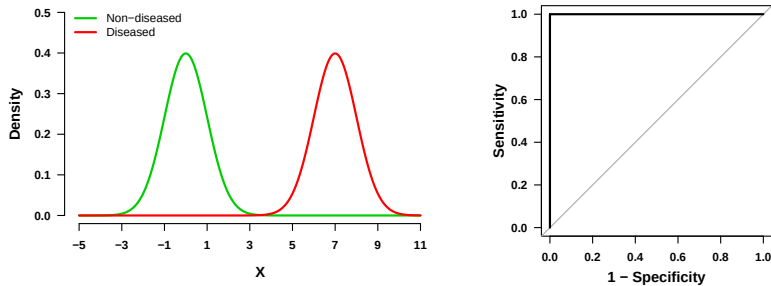


Figure 5: Perfect test

$AUC = 0.5 \rightarrow$ Chance line

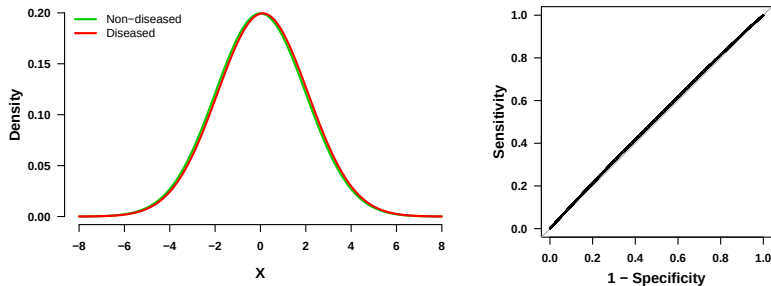


Figure 5: Uninformative test

In practice, diagnostic tests have ROC curves in between these two situations.

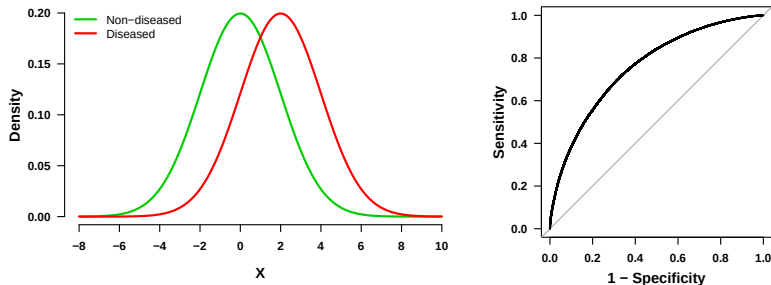


Figure 5: Common situation in real-world problems

Contents

- 1 Introduction
- 2 Continuous tests
- 3 ROC curves
- 4 Optimum threshold estimation**
- 5 Real-case example
- 6 Conclusions

Optimum threshold estimation

- Categorisation of continuous variables is sometimes disfavoured → loss of information [Cohen, 1983].
- However, it simplifies the classification of patients according to their state.
- Mean, median, quartiles, etc., are frequently used as thresholds, but they are arbitrary and highly sample-dependent.

Optimum threshold estimation

- Categorisation of continuous variables is sometimes disfavoured → loss of information [Cohen, 1983].
- However, it simplifies the classification of patients according to their state.
- Mean, median, quartiles, etc., are frequently used as thresholds, but they are arbitrary and highly sample-dependent.

Statistically viable strategies are needed for threshold estimation.

Optimum threshold estimation

- Aim: find a threshold that is optimal in some sense.
- The choice of a threshold depends on the trade-off that is considered to be acceptable between the two types of error (FP/FN).
- Plenty of methods have been proposed in the literature.

Youden index method

- Youden index maximisation [Youden, 1950]: c such that

$$J(c) = \text{Sens}(c) + \text{Spec}(c) - 1$$

is maximum.

- Threshold that leads to the point of the ROC curve farthest from the chance diagonal.

Youden index method

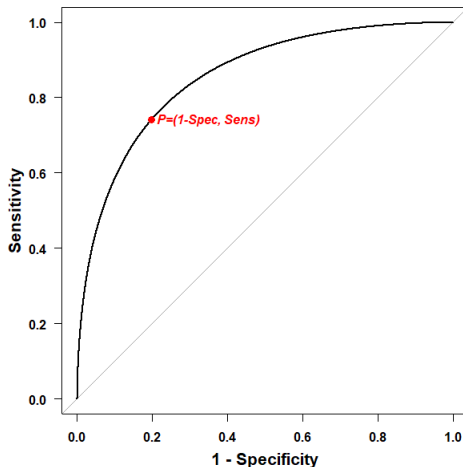


Figure 6: Illustration of Youden index maximisation

Youden index method

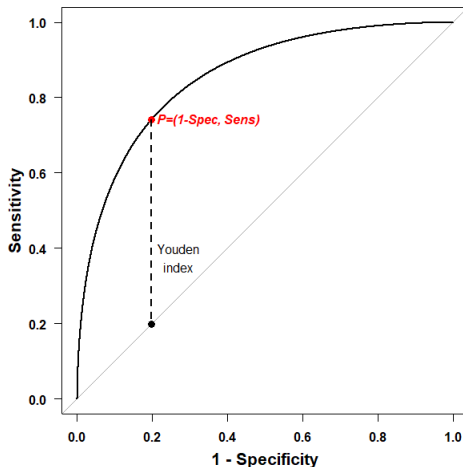


Figure 6: Illustration of Youden index maximisation

Distance to the ideal marker

- Threshold that leads to the point of the ROC curve closest to $(0, 1)$.
- c such that

$$\sqrt{(1 - \text{Spec}(c))^2 + (1 - \text{Sens}(c))^2}$$

is minimum.

Distance to the ideal marker

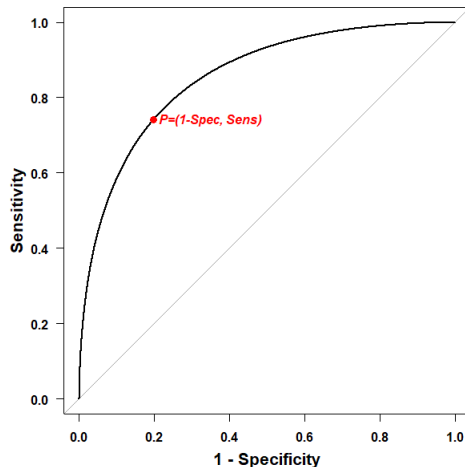


Figure 7: Illustration of minimisation of the distance to the ideal marker

Distance to the ideal marker

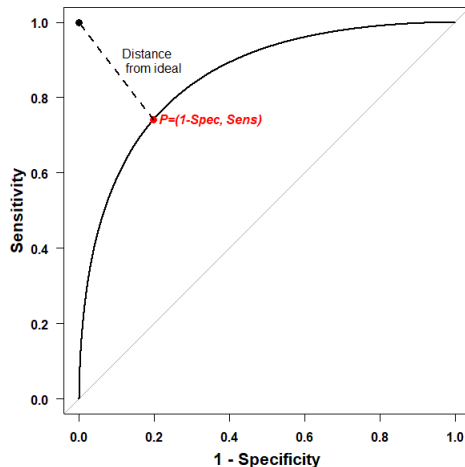


Figure 7: Illustration of minimisation of the distance to the ideal marker

Cost function approach

- A more general approach is to consider a cost function [Jund et al., 2005, Skaltsa et al., 2010].
- It accounts for all possible classifications: TP, TN, FP, FN.
- Cost function:

$$C(c) = TP(c) \cdot C_{TP} + FN(c) \cdot C_{FN} + FP(c) \cdot C_{FP} + TN(c) \cdot C_{TN},$$

where

C_{TP} : cost of a true positive,

C_{TN} : cost of a true negative,

C_{FP} : cost of a false positive,

C_{FN} : cost of a false negative.

Cost function based optimum threshold

The cost function $C(c)$ can be rewritten as:

$$C(c) = D(\text{Sens}(c) + R \cdot \text{Spec}(c) + G),$$

where

$$D = n\rho(C_{\text{TP}} - C_{\text{FN}}), \quad R = \frac{(1 - \rho)(C_{\text{TN}} - C_{\text{FP}})}{\rho(C_{\text{TP}} - C_{\text{FN}})},$$

$$G = \frac{\rho C_{\text{FN}} + (1 - \rho) C_{\text{FP}}}{\rho(C_{\text{TP}} - C_{\text{FN}})},$$

ρ = disease prevalence, n = sample size.

Cost function based optimum threshold

It can be shown that the cost minimising threshold c is the one such that

$$\frac{f_D(c)}{f_{\bar{D}}(c)} = R,$$

where $f_D(c)$ and $f_{\bar{D}}(c)$ are the density functions for the diseased and the non-diseased subpopulations, and

$$R = \left(\frac{1 - \rho}{\rho} \right) \cdot \left(\frac{C_{TN} - C_{FP}}{C_{TP} - C_{FN}} \right).$$

Cost function based optimum threshold

- This approach requires specification of costs for all classification results.
- C_{FP} , C_{FN} are allowed to be different → Flexibility and adaptation to the problem.
- $R = 1$ → Youden index maximisation.

Cost function based optimum threshold

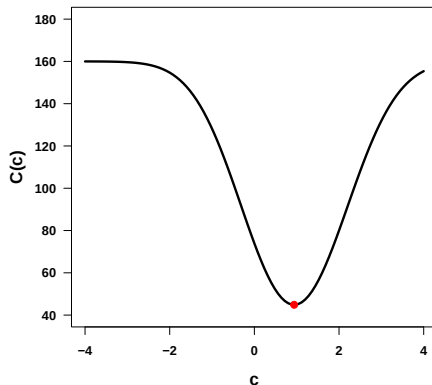


Figure 8: Example of cost function

Cost function based optimum threshold

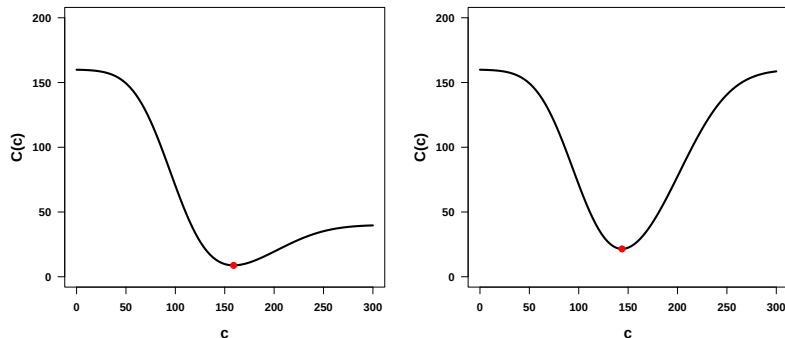


Figure 9: Left: $C_{TP} = C_{TN} = 0$, $C_{FP} = C_{FN} = 1$; right: $C_{TP} = C_{TN} = 0$, $C_{FP} = 1$, $C_{FN} = 4$

Estimation methods

- Parametric methods: binormality assumption.
- Non-parametric methods: empirical, smoothing.
- Confidence intervals: delta method, bootstrap.

Contents

- 1 Introduction
- 2 Continuous tests
- 3 ROC curves
- 4 Optimum threshold estimation
- 5 Real-case example**
- 6 Conclusions

Example: Prostate cancer data

- A study was conducted to evaluate the clinical usefulness of the level of acid phosphatase in blood serum (APBS) to predict nodal involvement.
- Data from $n = 53$ men diagnosed with prostate cancer [López-Ratón et al., 2017].
- Status: Nodal involvement present versus absent.
- Diagnostic test: APBS, in U/L·100.
- Nodal involvement present in 20 cases (37.7%).

Example: Prostate cancer data

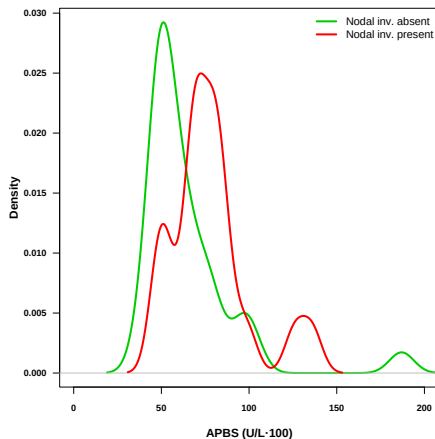


Figure 10: Prostate cancer data: density plot by status

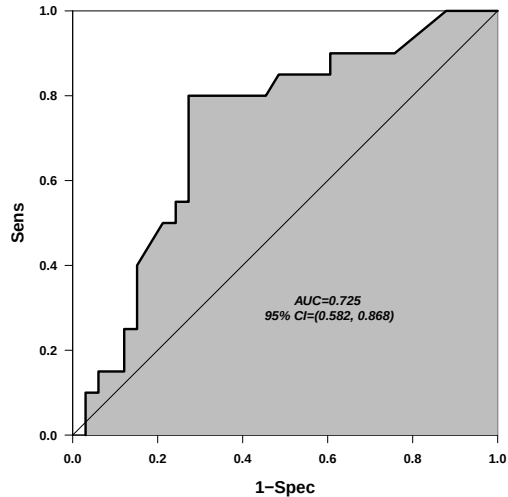


Figure 11: Prostate cancer data: ROC curve and AUC

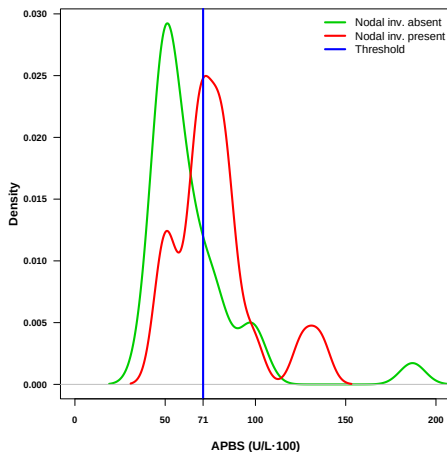


Figure 12: Prostate cancer data: densities and optimal threshold using the cost function approach (Youden index, $C_{TP} = C_{TN} = 0$, $C_{FP} = 1$, $C_{FN} = 1.65$)

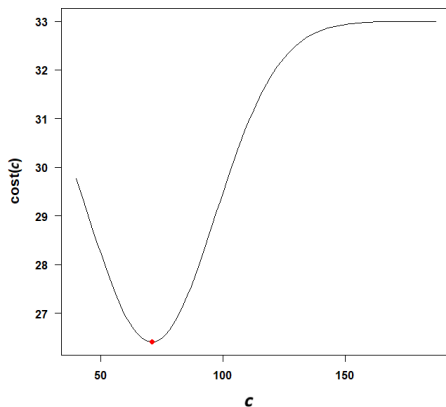


Figure 13: Prostate cancer data: Cost function assuming binormality (Youden index)

Software

R packages:

- `ThresholdROC` [Perez-Jaume et al., 2017]
- `GsymPoint` [López-Ratón et al., 2017]
- `cutpointr` [Thiele and Hirschfeld, 2021]

Contents

- 1 Introduction
- 2 Continuous tests
- 3 ROC curves
- 4 Optimum threshold estimation
- 5 Real-case example
- 6 Conclusions**

Conclusions

- Continuous medical tests are often used in diagnostic studies.
- ROC curves are a useful tool to evaluate their diagnostic ability.
- Youden index method is the most widely used for threshold estimation → Disadvantage: FP and FN are assumed to be equally undesirable.
- Multiple methods for optimum threshold estimation → Choice based on what “optimum” means in your problem.
- Related topics: k -state ($k > 2$) and survival problem.

Thank you for your attention!

References I



Cohen, J. (1983).

The cost of dichotomization.

Applied Psychological Measurement, 7(3):249–253.



Jund, J., Rabilloud, M., Wallon, M., and Ecochard, R. (2005).

Methods to estimate the optimal threshold for normally or log-normally distributed biological tests.

Medical Decision Making, 25(4):406–415.

PMID: 16061892.



López-Ratón, M., Molanes-López, E. M., Letón, E., and Cadarso-Suárez, C. (2017).

GsymPoint: An R package to estimate the generalized symmetry point, an optimal cut-off point for binary classification in continuous diagnostic tests.

The R Journal, 9:262–283.

<https://doi.org/10.32614/RJ-2017-015>.



Nakas, C. T., Bantis, L., and Gatsonis, C. (2023).

ROC Analysis for Classification and Prediction in Practice.

Chapman & Hall/CRC.

References II



Pepe, M. S. (2003).

The statistical Evaluation of Medical Tests for Classification and Prediction.
Oxford University Press, Oxford.



Perez-Jaume, S., Skaltsa, K., Pallarès, N., and Carrasco, J. L. (2017).

ThresholdROC: Optimum threshold estimation tools for continuous diagnostic tests in R.

Journal of Statistical Software, 82(4):1–21.



Skaltsa, K., Jover, L., and Carrasco, J. L. (2010).

Estimation of the diagnostic threshold accounting for decision costs and sampling uncertainty.

Biometrical Journal, 52(5):676–697.



Thiele, C. and Hirschfeld, G. (2021).

cutpointr: Improved estimation and validation of optimal cutpoints in R.

Journal of Statistical Software, 98(11):1–27.



Youden, W. J. (1950).

Index for rating diagnostic tests.

Cancer, 3(1):32–35.

References III



Zhou, X. H., Obuchowski, N. A., and McClish, D. K. (2011).
Statistical Methods in Diagnostic Medicine.
John Wiley & Sons, Hoboken, New Jersey.