

IDENTIFICATION OF BIOMARKERS OF ADVANCED PROSTATE CANCER USING MACHINE-LEARNING TOOLS



Author: Arnold, R. / Tutor: Kharashvili G, Dr., PhD.
Department of Clinical and Molecular Pathology



Lékařská fakulta

Introduction

There are several challenges in the prognosis of prostate cancer owing to its histologic and genetic heterogeneity. Nowadays, there are few accurate tissue-based biomarkers for prostate cancer aggressiveness. Intelligent processing of accumulated knowledge in the era of artificial intelligence is very promising. Machine learning is a subset of artificial intelligence. Its characteristics could be especially helpful in the management of prostate cancer, especially in digital pathology studies.

Objectives

To investigate the role of decision-support applications of machine learning in prostate cancer biomarker identification. We thought to evaluate the utility of the identified biomarkers and their combinations by developing a highly specific computer-aided algorithm.

Material and methods

Formalin fixed paraffin embedded tissues of 101 prostate carcinomas were stained immunohistochemically for Skp2, Slug, Ki67, p53, AR, PSA, E-cadherin, beta-catenin, CD151, vimentin, periostin and versican, and scored. Carcinomas were classified into localized, advanced and metastatic groups, and ISUP gleason grade groups. (Figure 1). Statistical analysis was performed by SPSS.

RULES FOR F (FALSE)

- Rule 1 for F**
if Ki67= low
and Ecadherin <= 75
and Ki67 <= 7
then F
- Rule 2 for F**
if Ki67 = low
and Ecadherin <= 75
and PSAtotal > 13,600
and beta catnucl = low
then F
- Rule 3 for F**
if Ki67= low
and Ecadherin > 75
and periostin cyt = low
and beta cat cyt <= 20
then F
- Rule 4 for F**
if Ki67= low
and Ecadherin > 75
and periostin cyt = low
and beta cat cyt >20and<= 35
and Skp2str = 0
then F
- Rule 5 for F**
if Ki67= low
and Ecadherin > 75
and periostin cyt= low
and beta cat cyt > 35
then F
- Rule 6 for F**
if Ki67 = low
and Ecadherin > 75
and periostin cyt = high
then F

Rule 7 for F

if Ki67= high
and Ecadherin = low
and Skp2str > 0
then F

Rule 8 for F

if Ki67= high
and Ecadherin = high
then F

RULES FOR T (TRUE)

Rule 1 for T
if Ecadherin <= 75
and Ki67> 7
and PSAtotal <= 13,600
then T

Rule 2 for T
if Ecadherin <= 75
and Ki67> 7
and PSAtotal > 13,600
and betacatnucl = high
then T

Rule 3 for T
if Ki67= low
and Ecadherin > 75
and betacat cyt > 20and<= 35
and Skp2str > 0
then T

Rule 4 for T
if Ki67 = high
and Ecadherin= low
and Skp2str <= 0
then T

Results

After biomarker evaluation and scoring, conditional target variable composited by patients cases with metastatic stage and high gleason grade group was created. Further, a decision tree (Figure 3) was developed with sets of characteristic rules (Figure 2) that display decision algorithms with their possible consequences, including events with a metastatic stage combined with a high GSgroup e.g. "true (T)" event - if both of these conditions are true and the "false (F)" event without either of these conditions. ROC curve and logistic regression models showed E-cadherin and Ki-67 as independent predictors with significance values of 0,018 and 0,048, respectively. (Figures 4a,b).

Conclusion

We were first to show possible use of machine learning algorithms in prediction of advanced prostate cancer, based on validated and novel tissue biomarkers.

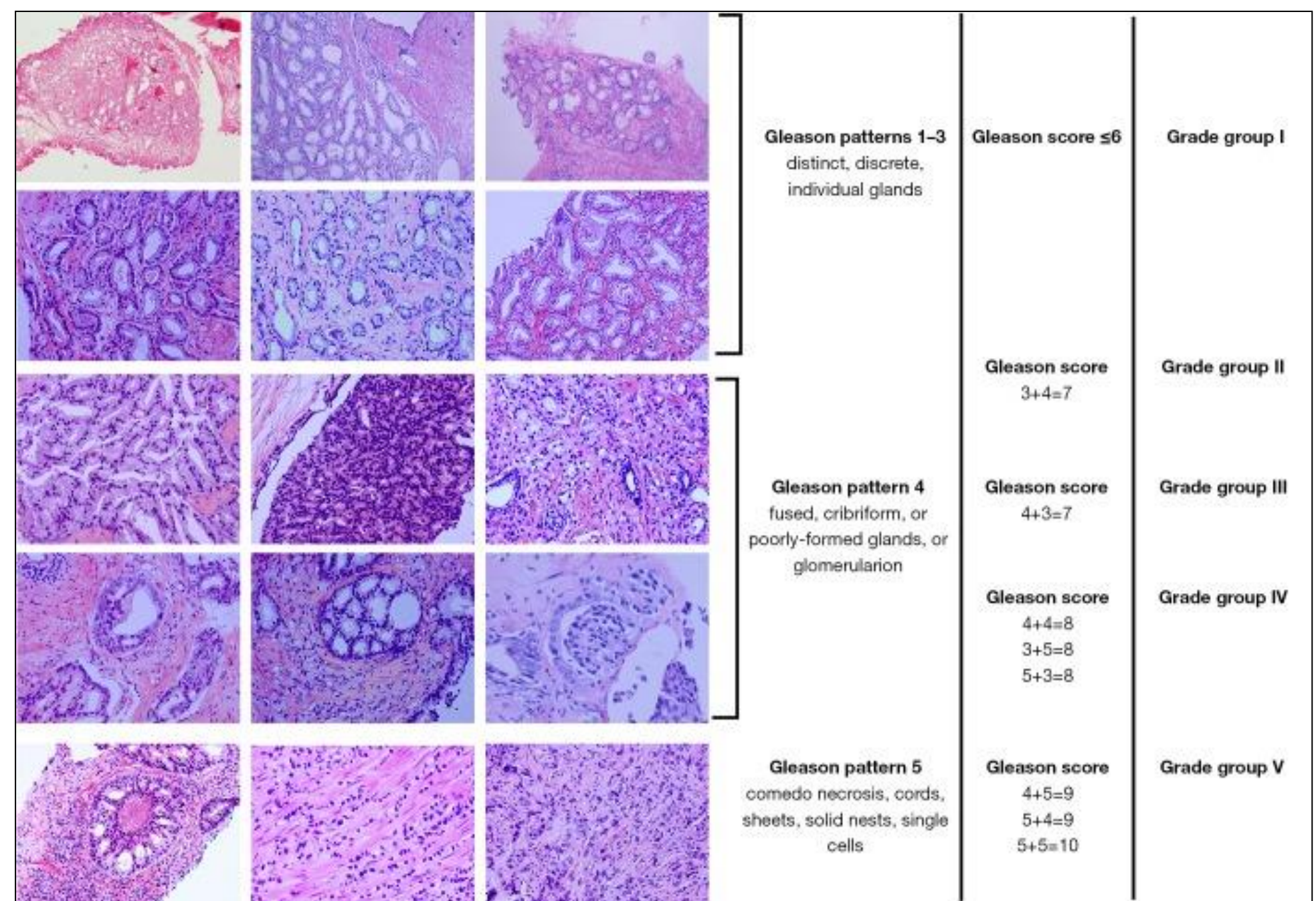


Figure 1. ISUP 2014 Gleason grade groups.

Figure 2. Decision algorithm rules for "true (T)" event - if both conditions (metastatic stage and high gleason group) are true and the "false (F)" event without either of these conditions.

References

- Epstein JI et al., Am J Surg Pathol. 2016;40(2):244-52.
- Mickova A, Kharashvili G et al. Int J Mol Sci. 2021, 22(6):2844.
- Jang WS et al.,Prostate Cancer Prostatic Dis. 2017;20(1):93-98.
- Baras AS et al.,Hum Pathol.2017;63:27-32.
- Trock BJ et al.,J Urol. 2009;182(4):1364-70.
- Zhao SG et al., JAMA Oncol. 2017;3(12):1663-1672.
- Prencipe M et al.,Prostate.2018 Apr 2.
- González-Reyes S et al.,Cancer Immunol Immunother. 2011;60(2):217-26.

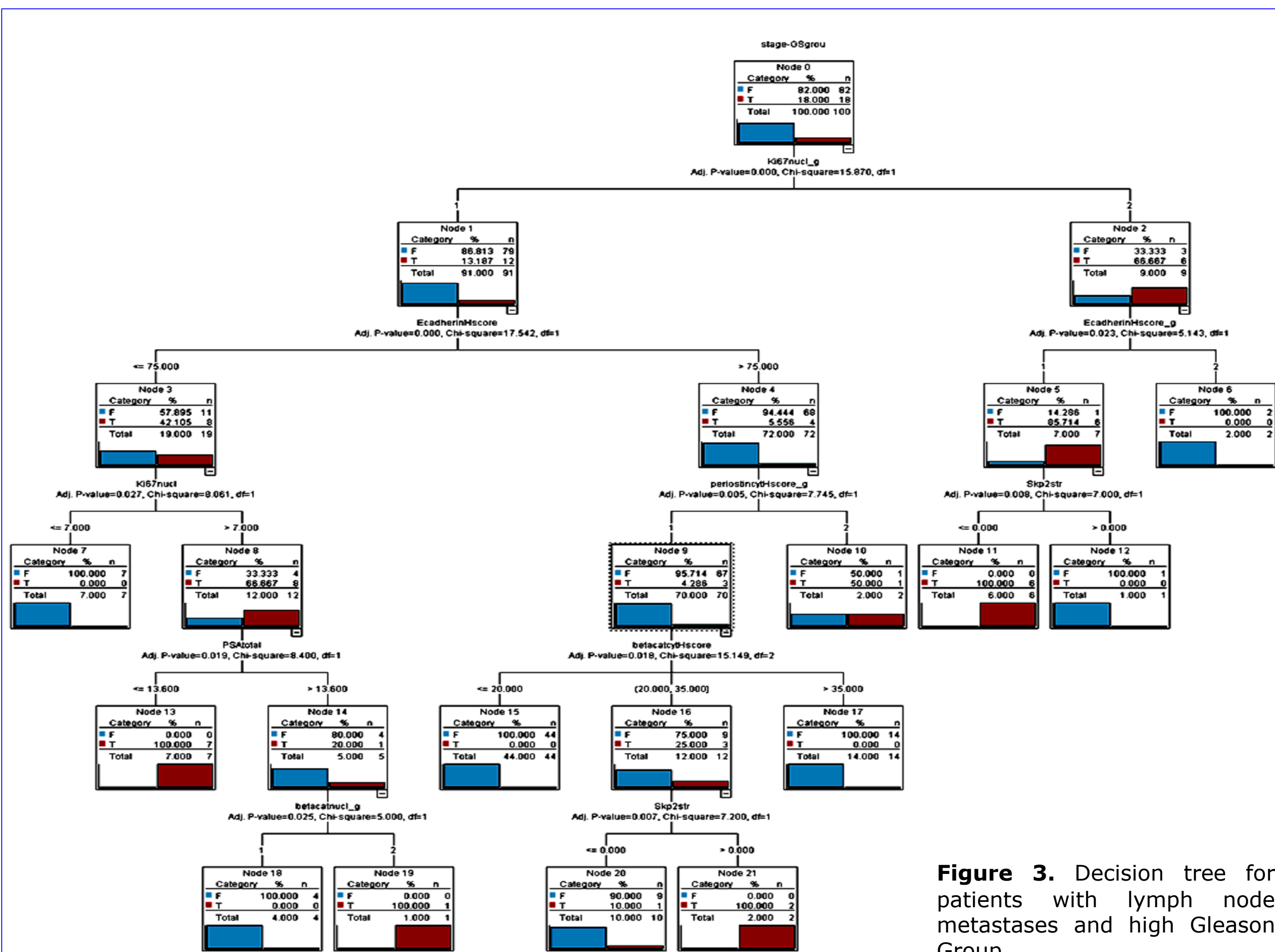


Figure 3. Decision tree for patients with lymph node metastases and high Gleason Group.

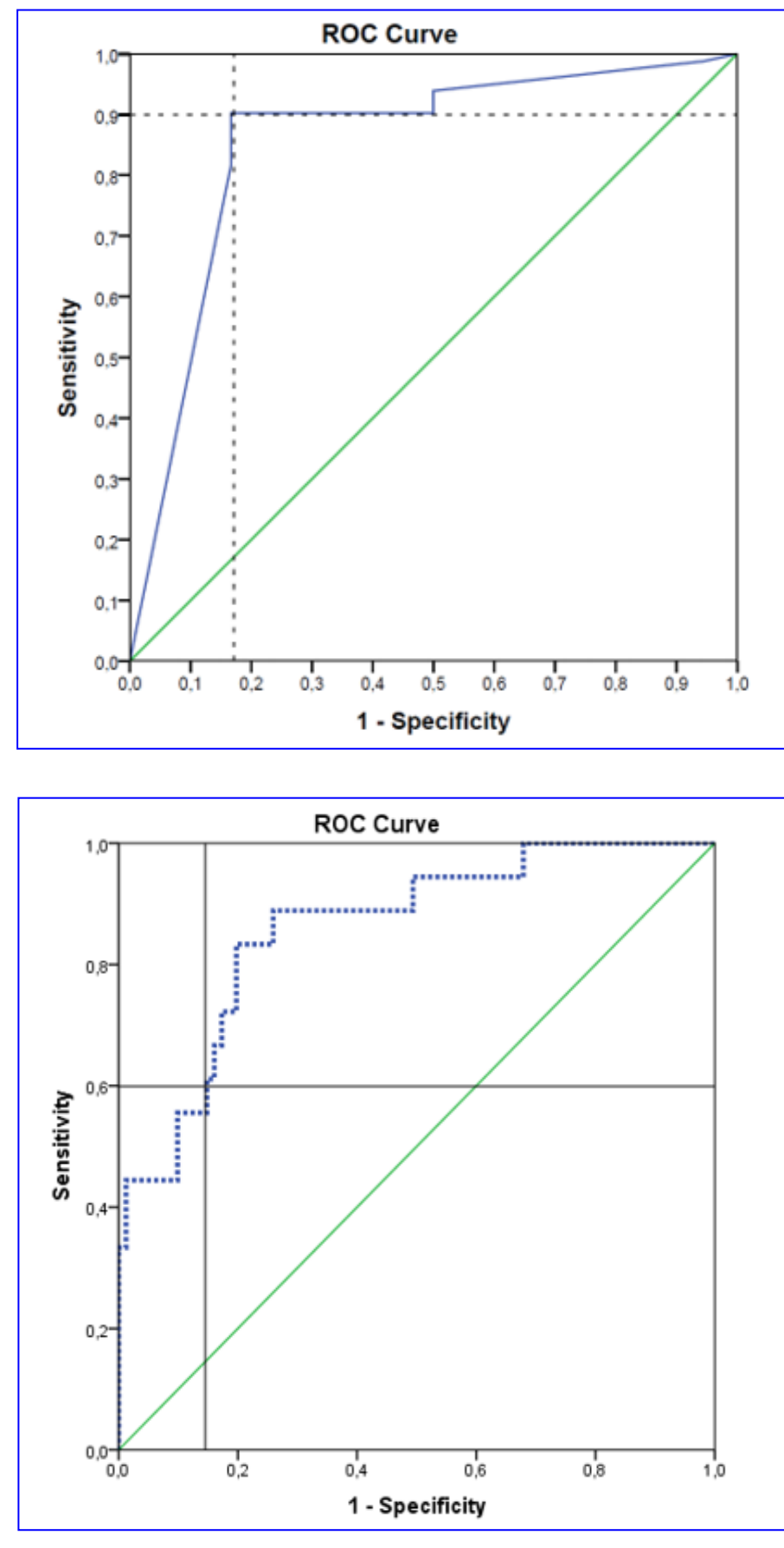


Figure 4. a. Roc curve for decision tree model. b. Roc curve for logistic regression model with the same variations as in decision tree. Area under curve 0,852 and 0,6, respectively.