Journal of Theoretical and Applied Information Technology
30th September 2022. Vol.100. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MACHINE LEARNING PREDICTION OF HEPATIC FIBROSIS IN HEPATITIS B EGYPTIAN PATIENTS BASED ON CLINICAL LABORATORY PARAMETERS

ESLAM TAHER SHARSHAR[1], HUDA AMIN MAGHAWRY[2], EMAN ABDELSAMEEA[3], NAGWA BADR[4].

[1,2,4]Information Systems Department, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt

[3]Hepatology and Gastroenterology Department, National Liver Institute, Menoufia University, Menoufia, Egypt

Email: [1]eslam.sharshar@cis.asu.edu, [2]huda_amin@cis.asu.edu.eg, [3]dreman555@yahoo.co.uk , [4]dr.nagwabadr@gmail.com

ABSTRACT

Liver fibrosis stage prediction in chronic hepatitis B virus (HBV) infected patients is vital. Liver biopsy is the reference style and gold standard to evaluate fibrosis stage but with many drawbacks. Therefore, using noninvasive methods are better alternatives. In this study, seven clinical laboratory parameters of 235 chronic HBV Egyptian patients with Hepatitis B virus were collected from HBV clinic at National Liver Institute that belongs to Menoufia University in Egypt. The aim of this study is to experiment multiple machine-learning methods based on clinical parameters to build efficient classification models that predict two liver related issues: the fibrosis stage and cirrhosis of liver in chronic HBV Egyptian patients. Also, attribute selection methods were applied to reduce the dimensionality and find the most relevant parameters. For fibrosis stage prediction, a classification model based on Logistic Regression achieved AUROC of 0.991 and accuracy of 93.61%. Besides, using only four parameters selected as the most relevant, AUROC of 0.971 and accuracy of 95.74% were achieved. For cirrhosis of liver prediction, a classification model based on Logistic Regression and cost sensitive with penalty of 2 achieved AUROC of 0.936 and accuracy of 91.49%. Besides, using only three parameters selected as the most relevant, AUROC of 0.92 and accuracy of 85.11% were achieved. The classification models outperformed noninvasive index-based method, FIB-4 that depends on four clinical parameters, in both fibrosis stage and liver cirrhosis prediction in chronic HBV Egyptian patients.

Keywords: *Machine-Learning; Attribute Selection; Chronic HBV; Fibrosis; Cirrhosis; FIB-4.*

1. INTRODUCTION

Hepatitis is a liver tissue inflammation. Hepatitis B virus (HBV) is one of the reasons of hepatitis. It can lead to hepatic fibrosis, cirrhosis and sometimes to hepatocellular carcinoma [1]. According to [2] more than 350 million people in the world are chronically injured with HBV. Early accurate assessment of hepatic fibrosis in chronic HBV patients is essential to start treatment early with appropriate therapy.

For many reasons, the liver biopsy is the standard way of measuring liver fibrosis. Firstly, it has superior quality and accuracy. Secondly, it has a standardized interpretation. Thirdly, it serves as a reference standard, and it is widely accepted [3]. However, it has a lot of limitations. It is invasive, susceptive to error in sampling, costly and difficult to be repeated [4-6]. Due to these limitations, the noninvasive alternative methods are very important to avoid the drawbacks of liver biopsy. These noninvasive methods have many advantages. They are safe, easy to perform, inexpensive and can be repeated.

There are two types of noninvasive methods. The first type is based on clinical parameters and blood tests such as fibrosis-4 index (FIB-4), the aspartate aminotransferase to platelet ratio index (APRI) and Fibrotest [7-9]. Some examples of clinical parameters and blood tests that used in this type are alanine aminotransferase (ALT), aspartate aminotransferase (AST) and platelet count (PLT). The second type is depending on techniques of liver imaging such as transient elastography [10][11].

The purpose of this study is to build an efficient classification model using machine learning methods to predict the fibrosis stage and cirrhosis of liver in chronic HBV Egyptian patients based on clinical laboratory parameters. Multiple machine learning were experimented to differentiate moderate from advanced fibrosis. Besides, they were experiemented to differentiate between non-cirrhosis and cirrhosis stage. A preprocessing steps were applied to the dataset such as removing almost empty patient records and dealing with missing values. In addition, attribute selection methods were applied to reduce the dimensionality and find the most relevant parameters. Then, machine learning methods were experimented using only those most relevant parameters. The results were compared to an index-based noninvasive method. The following sections of the paper are organized as follows: Section 2 introduces the related work, Section 3 presents material and methods, Section 4 introduces experiments and results, and finally the discussion and conclusion are stated in section 5 and 6, respectively.

2. RELATED WORK

In recent years, machine learning has been used in medical field for classification, prediction and diagnosis such as liver fibrosis stage prediction [12-17]. Many of researchers used machine learning methods to predict the fibrosis stages of liver and to discriminate between them in hepatitis B virus infected patients. Wang et al. [12] used three layers of Neural Network (Bayesian learning) based on serum and routine markers, such as age, AST, PLT and gamma-glutamyltransferase (GGT). They were the most important factors in the predictive model, to predict the significant fibrosis stages in chronic hepatitis B (CHB) patients. They achieved AUROC in training, validation and testing of 0.883, 0.884 and 0.92, respectively. Cao et al. [13] used seven clinical parameters, including age, ALT, AST, PLT, prothrombin time (PT), hemoglobin (HGB), and red cell distribution width (RDW), for building MLP classifier to discriminate between liver cirrhosis (LC) and non-LC cohort. They achieved an overall accuracy of 89.9% on test dataset. Wei et al. [16] used gradient boosting (GB) with four (age, ALT, AST and PLT) and six (age, ALT, AST, PLT, ALB and GGT) parameters to discriminate between early and advanced fibrosis and between fibrosis and cirrhosis. They achieved results better than FIB-4. Naiping et al. [17] established four machine learning methods (Random Forest Classifier (RFC), Decision Tree Classifier (DTC), Logistic Regression Classifier (LRC) and Support Vector Classifier (SVC)) using serum markers for liver fibrosis severity assessment (≥F2, ≥F3, F4). The results show that the RFC with 9 parameters (age, ALT, AST, GGT, Cre, WBC, L, PLT and INR) is feasible to assess severity for liver fibrosis with accuracy more than 83%. Due to the limited accuracy results and the large number of parameters included in classification, this study aims to propose an efficient classification model based on machine learning to improve the prediction of liver fibrosis stages and cirrhosis in chronic HBV Egyptian patients based on the most relevant clinical parameters.

3. MATERIALS AND METHODS

3.1. Data Preparation and Preprocessing

An HBV dataset was collected from HBV clinic at National Liver Institute that belongs to Menoufia University in Egypt from June 2016 to December 2019. The dataset consists of 282 instances. As a data cleaning step, 38 patients who had other diseases such as chronic hepatitis C virus (HCV) infection or hepatocellular carcinoma (HCC) were excluded. Patients who consumed alcohol almost daily should be excluded, but Egyptians rarely consume alcohol. Besides, 9 patients that have almost empty data were excluded. Thus, the final number of patients was 235.

Seven clinical laboratory parameters (features) were recorded including age, sex, aspartate aminotransferase (AST), alanine aminotransferase (ALT), total leucocytes count (TLC), polymerase chain reaction (PCR) quantitative level for DNA HBV, platelet count (PLT).

There were some missing data in some parameters. Regarding AST, there were 2 missing values representing 0.85% of the dataset. There were 38 missing values in TLC parameter representing 16.2% of the dataset. Regarding PLT parameter, there were 31 missing values representing 13.2% of the dataset. Expectation-maximization (EM) imputation techniques were used to deal with missing values using SPSS software [18][19]. According to the METAVIR system, the stages of liver fibrosis are scaled from F0 to F4 [12][20]. In order to differentiate moderate from advanced fibrosis, (F0-F2) reflected mild or moderate fibrosis and (F3 and F4) reflected

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

advanced fibrosis. To differentiate any fibrosis stages from cirrhosis of liver, (F0-F3) were considered as fibrosis and (F4) were considered as cirrhosis.

The clinical characteristics of patients' data in this study for the differentiation between moderate and advanced fibrosis and for the differentiation between non-cirrhosis and cirrhosis of liver were summarized in Table 1 and Table 2, respectively. SPSS software version 23.0 (SPSS Inc., Chicago, IL, USA) was used for all statistical analysis on the dataset. The data were viewed as mean ± standard deviation (SD).

Table 1: Clinical characteristics of the patients' dataset for differentiation between moderate and advanced fibrosis

Item	Moderate Fibrosis	Advanced Fibrosis
Number of patients	166	69
Gender (male/female)	(95/71)	(46/23)
Age (years)	35 ± 8.5	46.4 ± 10.6
HBV DNA level (IU/ml)	9425648 ± 93309238	3371133 ± 15796682
AST (U/L)	29.4 ± 19.6	47.3 ± 31.7
ALT (U/L)	37.8 ± 35.8	52.6 ± 47.9
TLC (10^9/L)	6.8 ± 1.6	6.1 ± 1.7
PLT (10^9/L)	220.7 ± 42	180.1 ± 50.7

Table 2: Clinical characteristics of the patients' dataset for the differentiation of non-cirrhosis and cirrhosis of liver

Item	Non-cirrhosis	Cirrhosis
Number of patients	197	38
Gender (male/female)	(119/78)	(22/16)
Age (years)	36.4 ± 9.5	48.5 ± 9.8
HBV DNA level (IU/ml)	16472974 ± 90463091	5581317 ± 21084668
AST (U/L)	31.7 ± 22	50.1 ± 33.4
ALT (U/L)	39.2 ± 34.4	57.8 ± 60.4
TLC (10^9/L)	6.7 ± 1.7	6.3 ± 1.6
PLT (10^9/L)	214.1 ± 46.6	181.4 ± 48.6

3.2. Classification Models Construction and Attribute Selection

In the case of discriminating between mild or moderate and advanced fibrosis, three machine learning methods were experimented to build the classification model using the training dataset. The three methods were Random Forest (RF) [21][22], Logistic Regression (LR) [23] and the third one is combination of RF and Logistic, which is called Vote [24]. But in the case of discriminating between non-cirrhosis and cirrhosis the situation was different. That is because the ratio between the two classes is relatively large (83.8%: 16.2%) unlike the first case where the ratio between the two classes (moderate and advanced) is (70.6%: 29.4%). The dataset was considered imbalanced. Therefore, Cost Sensitive approach was used [25-28]. It depends on making penalty of wrong prediction of minority class more than of majority class based on Logistic Regression and Random Forest with different penalties.

Besides, two attribute selection methods were experimented to reduce the dimensionality and find the most relevant parameters. The first method is Information Gain (IG) Attribute Evaluation that assesses the value of an attribute by measuring the information gain in relation to the class [29], which is calculated using equation (1):

$$IG\ (Class, Attribute) = H\ (Class) - H\ (Class \mid Attribute) \quad (1)$$

Where: *H (Class)* is the entropy of the class and *H (Class | Attribute)* measures the entropy of attribute by contributing to class.

The second method is Gain Ratio (GR) Attribute Evaluation that assesses the value of an attribute by measuring the gain ratio in relation to the class [29], which is calculated using equation (2):

$$GR\ (Class, Attribute) = IG\ (Class, Attribute) / H\ (Attribute) \quad (2)$$

Where: *H (Class)* is the entropy of the attribute.

Each model was constructed again after applying each attribute selection method using the training dataset with only the selected as most relevant parameters. Then, the results were compared. The classification model with the highest result was compared with noninvasive method, FIB-4 [8] with two cutoff values. Equation (3) is used to calculate FIB-4:

$$FIB\text{-}4 = \frac{Age(years) * AST(U/L)}{PLT(10^9/L) * \sqrt{ALT(U/L)}} \quad (3)$$

Where *Age, AST, PLT,* and *ALT* are clinical laboratory parameters.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. EXPERIMENTS AND RESULTS

For differentiation between moderate and advanced fibrosis, the dataset was divided into two classes. The first class was for mild or moderate stage (M) and consists of 166 patients representing 70.6% of the dataset. The second class was for advanced stage (A) and consists of 69 patients representing 29.4% of the dataset. For the differentiation of non-cirrhosis and cirrhosis of liver, the dataset was divided into two classes. The first class was for non-cirrhosis class (F), which consists of 197 patients representing 83.8% of the dataset. The second class was for cirrhosis class (C), which consists of 38 patients representing 16.2% of the dataset.

The dataset was split into 80% for training dataset and 20% for test dataset keeping the same ratio between the two classes. Therefore, in the case of fibrosis stage prediction, the training dataset consisted of 132 patients for mild or moderate class and 56 patients for advanced class. The test dataset consisted of 34 patients for mild or moderate class and 13 patients for advanced class. In the case of cirrhosis prediction, the training dataset consisted of 158 patients for non-cirrhosis class and 30 patients for cirrhosis class. The test dataset consisted of 39 non-cirrhosis patients and 8 cirrhosis patients.

Waikato Environment for Knowledge Analysis (WEKA) version (3.9.3) was used to build and evaluate each classification model with 10-fold cross-validation to avoid over fitting [13][30][31]. WEKA is also used to apply both attribute selection methods. The evaluation measures used to evaluate and compare the classification models are accuracy (ACC), sensitivity (SE) and area under receiver operating characteristic curves (AUROC).

4.1. Differentiation between Mild or Moderate Fibrosis Stage (F0-F2) and Advanced Fibrosis Stage (F3 And F4)

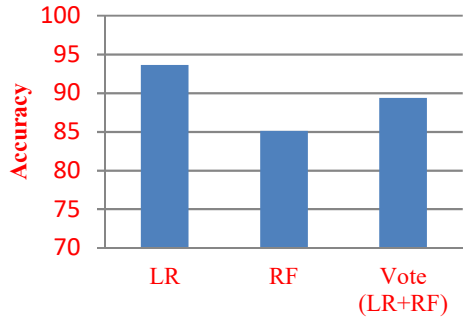
The training dataset was used to construct three models based on three supervised machine-learning methods: Random Forest (RF), Logistic Regression (LR) and Vote (RF+LR). Then, the models were evaluated using the test dataset to measure their performance. The results are shown in Table 3 and Figures 1 and 2. Using the test dataset, the model that achieved the highest results in differentiating between the two classes was Logistic Regression based classification model. It achieved AUROC of 0.991, accuracy of 93.61% and sensitivity of 0.936 (0.971 for moderate fibrosis class and 0.846 for advanced fibrosis class). This means that the best way to differentiate between the two classes was linear separation more than any other method.

Table 3: Results of fibrosis stage prediction using seven clinical laboratory parameters

Classifier	Fibrosis Stage	Accuracy	Sensitivity		AUROC
LR	**M**	**93.61%**	**0.971**	**0.936**	**0.991**
	A		**0.846**		
RF	M	85.11%	0.882	0.851	0.928
	A		0.769		
Vote (RF + LR)	M	89.36%	0.912	0.894	0.973
	A		0.846		

Journal of Theoretical and Applied Information Technology
30th September 2022. Vol.100. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



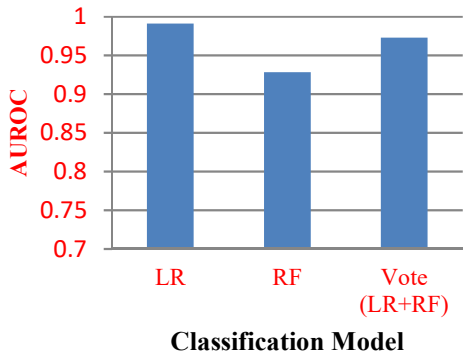


Figure 1: Comparison of classification models using seven clinical laboratory parameters in terms of accuracy for fibrosis stage prediction

Figure 2: Comparison of classification models using seven clinical laboratory parameters in terms of AUROC for fibrosis stage prediction

Then, two attribute selection methods were applied: Information Gain (IG) Attribute Evaluation and Gain Ratio (GR) Attribute Evaluation on training dataset. According to the relation values of all parameters with class, the number of parameters that showed the best relationship with class was reduced to four: Age, AST, ALT and PLT, as shown in Table 4.

Table 4: The relation values of the seven parameters with class in case of fibrosis stage prediction

Parameter	Gain Ratio (GR) Attribute Evaluation	Information Gain (IG) Attribute Evaluation
AST	**0.20425**	**0.14614**
Age	**0.14814**	**0.13296**
ALT	**0.10796**	**0.08475**
PLT	**0.10321**	**0.10267**
Sex	0.00475	0.0046
HBV DNA level	0	0
TLC	0	0

The three models were constructed with only the four parameters. The results were shown in Table 5 and Figures 3 and 4. Using only four parameters, the highest results achieved by Logistic Regression based classification model as with using all parameters as expected. It achieved AUROC of 0.971, accuracy of 95.74% and sensitivity of 0.957 (1 for moderate fibrosis class and 0.846 for advanced fibrosis class) However, the classification model achieved better results using the seven clinical parameters.

Table 5: Results of fibrosis stage prediction using the selected four clinical laboratory parameters (Age, AST, ALT and PLT)

Classifier	Fibrosis Stage	Accuracy	Sensitivity		AUROC
LR	M	**95.74%**	**1**	**0.957**	**0.971**
	A		**0.846**		
RF	M	85.11%	0.882	0.851	0.854
	A		0.769		
Vote (RF + LR)	M	89.36%	0.941	0.894	0.934
	A		0.769		

Journal of Theoretical and Applied Information Technology
30th September 2022. Vol.100. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

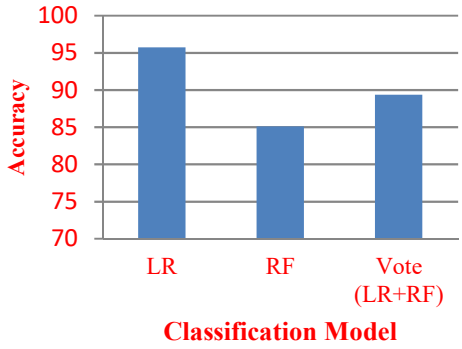


Figure 3: Comparison of classification models using four clinical laboratory parameters in terms of accuracy for fibrosis stage prediction

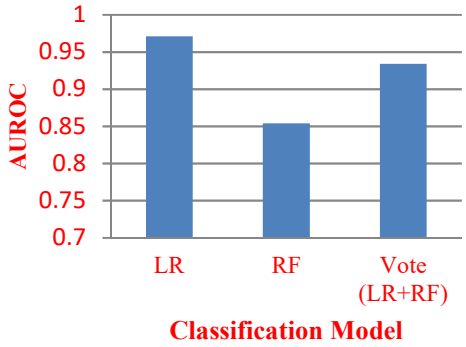


Figure 4: Comparison of classification models using four clinical laboratory parameters in terms of AUROC for fibrosis stage prediction

4.2. Differentiation between Non- Cirrhosis (F0-F3) and Cirrhosis (F4) of Liver

As mentioned before, the ratio between the two classes in the training dataset is relatively large (83.8%: 16.2%). Therefore, Cost Sensitive (CS) Classifier was used with three different penalty values (2,3 and 5). Two machine learning methods were applied: Logistic Regression (LR) and Random Forest (RF). The models were evaluated using the test dataset. The results are shown in Table 6 and Figures 5 and 6. In this study, the Cost Sensitive Classifier model with Logistic Regression using penalty value of 2 achieved the best results (AUROC = 0.936, accuracy = 91.49% and sensitivity = 0.915 (0.923 for non-cirrhosis class and 0.875 for cirrhosis class)). This means that the best way to differentiate between the mild or moderate and advanced fibrosis was linear separation more than any other method.

Table 6: Results of liver cirrhosis prediction using seven clinical laboratory parameters

Classifier	Fibrosis Stage	Accuracy	Sensitivity		AUROC
LR	F	85.11%	0.923	0.851	0.936
	C		0.500		
LR and CS (2)	F	91.49%	0.923	0.915	0.936
	C		0.875		
LR and CS (3)	F	87.23%	0.846	0.872	0.923
	C		1.000		
LR and CS (5)	F	76.59%	0.718	0.766	0.917
	C		1.000		
RF	F	82.98%	0.949	0.830	0.893
	C		0.250		
RF and CS (2)	F	87.23%	0.923	0.872	0.888
	C		0.625		
RF and CS (3)	F	89.36%	0.897	0.894	0.891
	C		0.875		
RF and CS (5)	F	76.59%	0.744	0.766	0.886
	C		0.875		

Journal of Theoretical and Applied Information Technology
30th September 2022. Vol.100. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

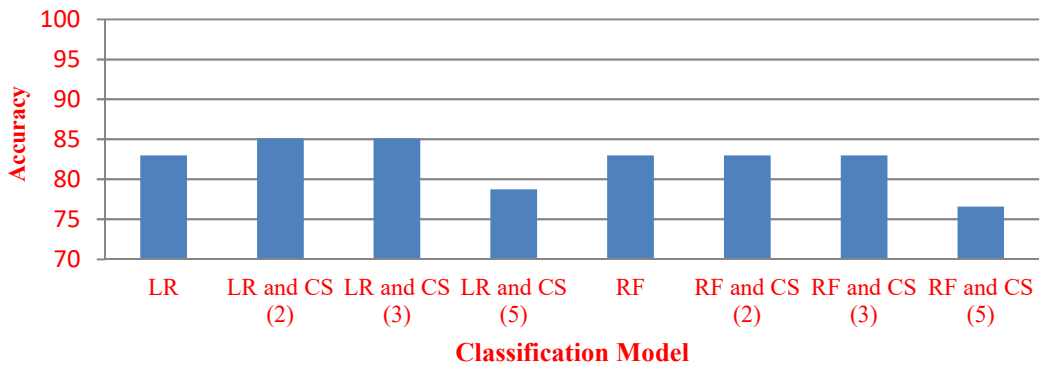


Figure 5: Comparison of classification models using seven clinical laboratory parameters in terms of accuracy for liver cirrhosis prediction

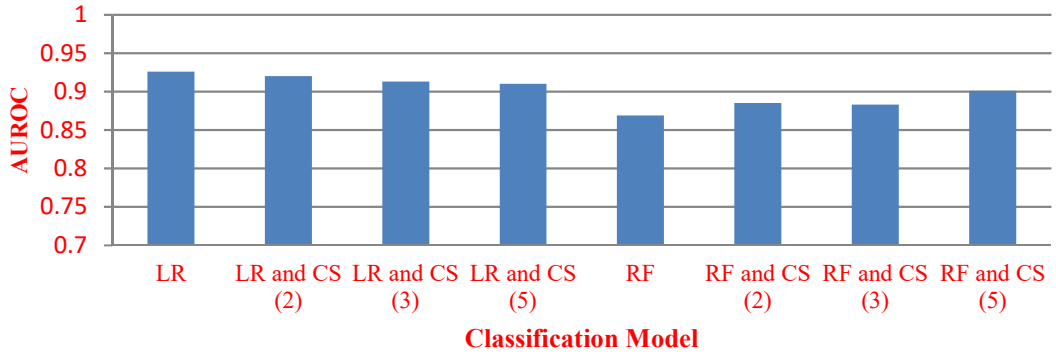


Figure 6: Comparison of classification models using seven clinical laboratory parameters in terms of AUROC for liver cirrhosis prediction

Then, the two attribute selection methods mentioned before were applied on the training dataset. According to the relation values of all parameters with class, the number of parameters that showed the best relationship with class was reduced to three: Age, AST, and PLT, as shown in Table 7.

Table 7: The relation values of the seven parameters with class in case of cirrhosis prediction

Parameter	Gain Ratio (GR) Attribute Evaluation	Information Gain (IG) Attribute Evaluation
AST	**0.119843**	**0.117632**
Age	**0.096109**	**0.092276**
PLT	**0.065158**	**0.064817**
Sex	0.000467	0.000445
ALT	0	0
HBV DNA level	0	0
TLC	0	0

The models were constructed with only the three parameters selected as the most relevant. The results are shown in Table 8 and Figures 7 and 8. Using only three parameters, the highest results were achieved by Cost Sensitive Classifier with Logistic Regression and penalty value of 2. It showed higher results than other classifiers using test dataset (AUROC = 0.92, accuracy = 85.11% and sensitivity = 0.851 (0.872 for non-cirrhosis class and 0.750 for cirrhosis class).

Table 8: Results of liver cirrhosis prediction using the selected three clinical laboratory parameters (Age, AST and PLT)

Classifier	Fibrosis Stage	Accuracy	Sensitivity		AUROC
LR	F	82.98%	0.923	0.83	0.926
	C		0.375		
LR and CS (2)	**F**	**85.11%**	**0.872**	**0.851**	**0.920**
	C		**0.750**		
LR and CS (3)	F	85.11%	0.821	0.851	0.913
	C		1.000		
LR and CS (5)	F	78.72%	0.744	0.787	0.910
	C		1.000		
RF	F	82.98%	0.923	0.830	0.869
	C		0.375		
RF and CS (2)	F	82.98%	0.872	0.830	0.885
	C		0.625		
RF and CS (3)	F	82.98%	0.821	0.830	0.883
	C		0.875		
RF and CS (5)	F	76.59%	0.744	0.766	0.901
	C		0.875		

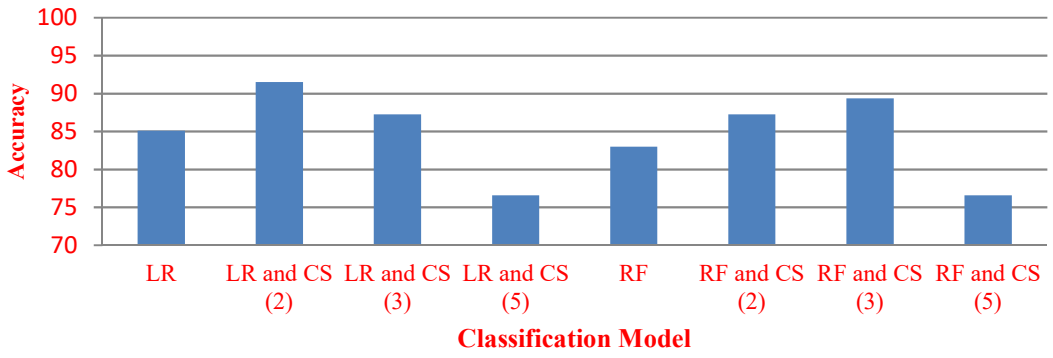


Figure 7: Comparison of classification models using three clinical laboratory parameters in terms of accuracy for liver cirrhosis prediction

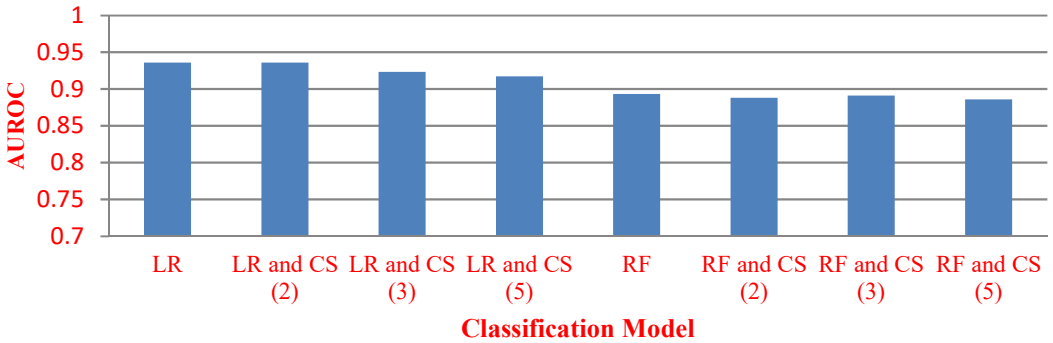


Figure 8: Comparison of classification models using three clinical laboratory parameters in terms of AUROC for liver cirrhosis prediction

Journal of Theoretical and Applied Information Technology
30th September 2022. Vol.100. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.3. Comparison with the Noninvasive Method FIB-4

Based on clinical laboratory parameters, the classification models with the highest achieved results were compared to FIB-4 method. Two cutoff values were used to prove the efficiency of the classification models.

For the discrimination between mild or moderate and advanced fibrosis stage, the Logistic Regression achieved the highest results. Therefore, it is compared with FIB-4 using two cutoff values. The first suggested cutoff value was 1.45. It is used in many previous researches [16][32-34]. The second value was 1.06. It was identified by Youden's Index [35-37] that depends on maximizing the sum of sensitivity and specificity in ROC. SPSS is used and dataset of 235 patients were used. The Logistic Regression classifier outperformed FIB-4 method as shown in Table 9 and Figures 9 and 10. It achieved AUROC of 0.991 and 0.971 using seven and four parameters, respectively.

Table 9: Comparison between Logistic Regression and FIB-4 with two cutoff values for predicting fibrosis stage

Classifiers	AUROC	Sensitivity	Accuracy
LR 7 Parameters	**0.991**	**0.936**	**93.61%**
LR 4 Parameters	**0.971**	**0.957**	**95.74%**
FIB-4 (1.45)	0.870	0.915	91.48%
FIB-4 (1.06)	0.912	0.872	87.23%

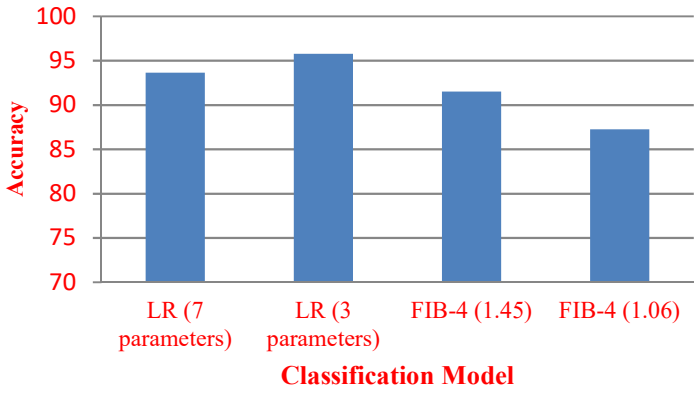


Figure 9: Comparison between Logistic Regression classifiers and FIB-4 with two cutoff values in terms of accuracy for fibrosis stage prediction

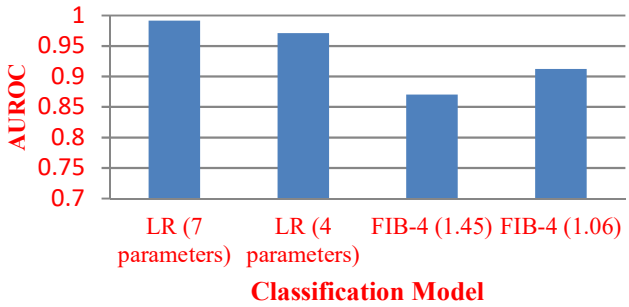


Figure 10: Comparison between Logistic Regression classifiers and FIB-4 with two cutoff values in terms of AUROC for fibrosis stage prediction

For the discrimination between non-cirrhosis and cirrhosis, the Cost Sensitive Classifier with penalty value of 2 with Logistic Regression achieved the highest results. Therefore, it is compared with FIB-4 using two cutoff values. The first cutoff value was 1.45 [16]. The second value was 1.13. It was identified by Youden's Index. The Cost Sensitive Classifier with Logistic Regression outperformed FIB-4 method as shown in Table 10 and Figures 11 and 12. It achieved AUROC of 0.936 and 0.920 using seven and three parameters, respectively.

Table 10: Comparison between Cost Sensitive Classifier with Logistic Regression and FIB-4 with two cutoff values for predicting liver cirrhosis

Classifiers	AUROC	Sensitivity	Accuracy
LR and CS (2) 7 Parameters	**0.936**	**0.915**	**91.49%**
LR and CS (2) 3 Parameters	**0.920**	**0.851**	**85.11%**
FIB-4 (1.45)	0.736	0.808	80.85%
FIB-4 (1.13)	0.796	0.745	74.47%

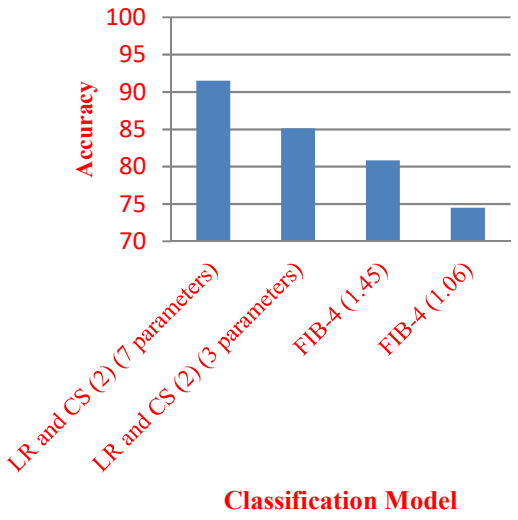


Figure 11: Comparison between Cost Sensitive Classifier with Logistic Regression and FIB-4 with two cutoff values in terms of accuracy for liver cirrhosis prediction

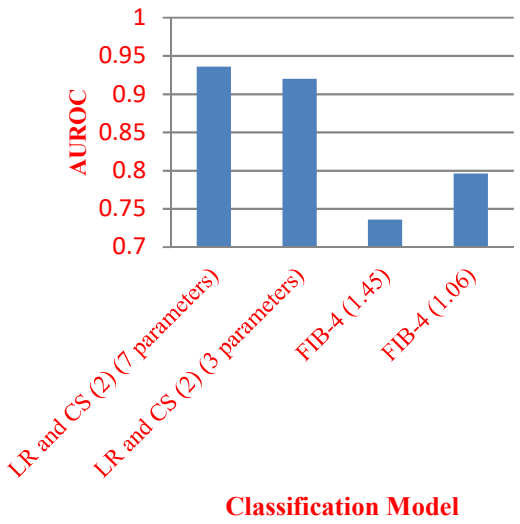


Figure 12: Comparison between Cost Sensitive Classifier with Logistic Regression and FIB-4 with two cutoff values in terms of AUROC for liver cirrhosis prediction

5. DISCUSION

In this study, supervised machine learning methods were experimented on chronic HBV infected Egyptians patients to predict fibrosis stages and cirrhosis. Three different classification models were constructed to differentiate between moderate and advanced fibrosis stages. The models were Logistic Regression, Random Forest and Vote. Besides, eight different classification models were constructed to differentiate between non-cirrhosis and cirrhosis of liver (Logistic Regression, Cost Sensitive Classifier with Logistic Regression using three different penalties, Random Forest and Cost Sensitive Classifier with Random Forest using three different penalties). Seven clinical parameters were used: Age, Sex, AST, ALT, HBV DNA, TLC and PLT. Two attributes selection methods were also experimented trying to improve the classification models: Information Gain (IG) Attribute Evaluation and Gain Ratio (GR) Attribute Evaluation.

Logistic Regression achieved the best results in differentiating between moderate and advanced fibrosis. The model was compared to noninvasive method, FIB-4 that depends on four clinical parameters using two cutoff values (1.45 and 1.06). Logistic Regression showed better performance in differentiation and prediction of the two classes (AUROC = 0.991) than FIB-4 (AUROC = 0.912 (cutoff =1.06) and AUROC = 0.870 (cutoff =1.45)). The classification models were trained using four parameters that showed the best relationship with class using attribute selection

methods: AST, Age, ALT and PLT. The accuracy of Random Forest and Vote models in both cases (with all parameters or with only selected parameters) were the same, but area under ROC was better with all parameters. Applying Logistic Regression with the selected four parameters achieved an accuracy of 95.74% that is better by 2.1% than when using all parameters which is 93.61%. But area under ROC using all parameters was 0.991, which is better than the model with only four selected parameters, which is 0.971. However, it outperformed noninvasive method, FIB-4, using the same number of parameters, which is four.

Logistic Regression with Cost Sensitive Classifier achieved the best results in differentiating between non-cirrhosis and cirrhosis. The model was compared to FIB-4 that depends on four clinical parameters using two cutoff values (1.45 and 1.13). It showed better performance in differentiation and prediction of the two classes with AUROC of 0.936 than FIB-4 with AUROC of 0.796 (cutoff =1.13) and AUROC = 0.736 (cutoff =1.45). No improvements were shown when the classifiers were applied using the parameters that showed the best relationship with class using attribute selection methods: Age, AST and PLT. However, using only the three parameters, Logistic Regression with Cost Sensitive Classifier outperformed FIB-4, which depends on four parameters with AUROC of 0.920.

6. CONCLUSION

Many viruses cause liver hepatitis. One of them is HBV. Many people in the world are infected with HBV. This infection causes fibrosis of liver which is graded from F0 to F4 depending on METAVIR system or in other words from mild or moderate fibrosis to advanced fibrosis and cirrhosis. The liver biopsy is a standard method for fibrosis stages evaluation but with many limitations. Therefore, the need for noninvasive alternative methods has become very important to avoid the disadvantages of biopsy as much as possible. Many of these methods were created depending on clinical parameters such as FIB-4 and APRI, and other based on liver imaging as transient elastography. In this study, based on seven clinical parameters of 235 chronic HBV Egyptian patients with Hepatitis B virus, efficient machine learning based classification models were proposed for the prediction of fibrosis stages (mild or moderate and advanced fibrosis), and cirrhosis of liver in HBV Egyptian patients to avoid the liver biopsy. An HBV dataset was collected. Data preparation and preprocessing were performed. Then, classification models were built using all clinical parameters in the dataset. Also attribute selection methods were experimented to reduce the number of parameters and to improve the performance of the classification models. The same classification models were built using only four parameters for differentiation between mild or moderate and advanced fibrosis stage. Besides, they were built using three parameters for differentiation between non-cirrhosis and cirrhosis stage. For the differentiation between mild or moderate and advanced fibrosis stage, AUROC of 0.991, accuracy of 93.61% and sensitivity of 0.936 (0.971 for moderate fibrosis class and 0.846 for advanced fibrosis class) were achieved using Logistic Regression. Using only four attributes (AST, Age, ALT and PLT), AUROC of 0.971 and accuracy of 95.74% were achieved. For the differentiation between non-cirrhosis and cirrhosis stage, AUROC of 0.936, accuracy of 91.49% and sensitivity of 0.915 (0.923 for non-cirrhosis class and 0.875 for cirrhosis class) were achieved using Cost Sensitive Classifier with logistic classifier with penalty value of 2. Using only three attributes (AST, Age, and PLT), AUROC of 0.920 and accuracy of 85.11% were achieved. The results of the classification models were compared to noninvasive index-based method, FIB-4 that depends on four clinical parameters. The classification models outperformed FIB-4 to avoid liver biopsy limitation using only four and three parameters for differentiation between mild or moderate and advanced fibrosis stage, and differentiation between non-cirrhosis and cirrhosis stage, respectively.

For future work, the classification models will be experimented on Egyptian patients' HBV data from different liver institutes as well as from different regions of world to avoid the drawbacks of liver biopsy and to discover liver fibrosis in the early stages.

7. ACKNOWLEDGMENT

We would like to thank and acknowledge the support that was given by Dr. Nashwa Shebl and Dr. Warda Othman, Hepatology and Gastroenterology Department, National Liver Institute, Menoufia University.

REFERENCES

[1] W. Maddrey, Hepatitis B: an important public health issue, J. Med. Virol. 61 (2000), pp. 362-366.

[2] Lavanchy, Hepatitis B virus epidemiology, disease burden, treatment, arid current and emerging prevention and control measures, *Journal of Viral Hepatitis* 11 (2) (2004), pp. 97–107.

[3] A.A. Bravo, S.G. Sheth, S. Chopra, Liver biopsy, N. Engl. J. Med. 344 (2001), pp. 495-500.

[4] L.B. Seeff, G.T. Everson, T. R. Morgan, T.M. Curto, W.M. Lee, M.G. Ghany, M.L. Shiffman, R.J. Fontana, A.M. Di Bisceglie, H.L. Bonkovsky, et al., Complication rate of percutaneous liver biopsies among persons with advanced chronic liver disease in the HALT-C trial, Clinical Gastroenterology and Hepatology 8 (10) (2010), pp. 877–883.

[5] A. Regev, M. Berho, L.J. Jeffers, C. Milikowski, E.G. Molina, N.T. Pyrsopoulos, Z.Z. Feng, K.R. Reddy, E.R. Schiff, Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection, The American Journal of Gastroenterology 97 (10) (2002), pp. 2614–2618.

[6] P. Thampanitchawong, T. Piratvisuth, Liver biopsy: complications and risk factors, World J. Gastroenterol. 5 (4) (1999), pp. 301–304.

[7] C.T. Wai, J.K. Greenson, R.J. Fontana, J.D. Kalbfleisch, J.A. Marrero, H.S. Conjeevaram, A.S. Lok, A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C, Hepatology 38 (2003), pp. 518-526.

[8] R.K. Sterling, E. Lissen, N. Clumeck, R. Sola, M.C. Correa, J. Montaner, M.S. Sulkowski, F.J. Torriani, D.T. Dieterich, D.L. Thomas, et al., Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection, *Hepatology* 43 (2006), pp. 1317-1325.

[9] P. Calès, F. Oberti, S. Michalak, I. Hubert-Fouchard, M.C. Rousselet, A. Konaté, Y. Gallois, C. Ternisien, A. Chevailler, F. Lunel, A novel panel of blood markers to assess the degree of liver fibrosis, *Hepatology* 42 (2005), pp. 1373-138.

[10] C.R.A. Lesmana, S. Salim, I. Hasan, A.S. Sulaiman, R.A. Gani, L.S. Pakasi, L.A. Lesmana, E. Krisnuhoni, U. Budihusodo, Diagnostic accuracy of transient elastography (FibroScan) versus the aspartate transaminase to platelet ratio index in assessing liver fibrosis in chronic hepatitis B: the role in primary care setting, Journal of Clinical Pathology 64 (10) (2011), pp. 916–920.

[11] L. Sandrin, B. Fourquet, J.M. Hasquenoph, S. Yon, C. Fournier, F. Mal, C. Christidis, M. Ziol, B. Poulet, F. Kazemi, et al., Transient elastography: a new noninvasive method for assessment of hepatic fibrosis, Ultrasound Med. Bi.ol. 29 (12) (2003), pp. 1705–1713.

[12] Wang, Q. Wang, F. Shan, B. Liu, and C. Lu, Identification of the risk for liver fibrosis on CHB patients using an artificial neural network based on routine and serum markers, *BMC Infectious Diseases* 10 (article 251) (2010).

[13] Y. Cao, Z.D. Hu, X.F. Liu, A.M. Deng, C.J. Hu, An MLP Classifier for Prediction of HBV-Induced Liver Cirrhosis Using Routinely Available Clinical Parameters, Disease Markers journal 35 (6) (2013), pp. 653–660.

[14] A.M. Hashem, M.E. Rasmy, K.M. Wahab, O.G. Shaker, Single stage and multistage classification models for the prediction of liver fibrosis degree in patients with chronic hepatitis C infection, Computer Methods and Programs in Biomedicine 105 (2012), pp. 194-209

[15] S. Hashem, G. Esmat, W. Elakel, S. Habashy, S. Abdel Raouf, S. Darweesh. M. Soliman, M. Elhefnawi, M. El-Adawy, M. ElHefnawi, Accurate Prediction of Advanced Liver Fibrosis Using the Decision Tree Learning Algorithm in Chronic Hepatitis C Egyptian Patients, Gastroenterology Research and Practice 2016 (article ID 2636390) 7 pages.

[16] R. Wei, J. Wang, X. Wang, G. Xi, Y. Wang, H. Zhang, C.Y. Peng, C. Rajani, S. Kwee, P. Liu, et al., Clinical prediction of HBV and HCV related hepatic fibrosis using machine learning, EBioMedicine 35 (2018) 124-132.

[17] N. Li, J. Zhang, S. Wang, Y. Jiang, J. Ma, J. Ma, L. Dong, G. Gong, Machine Learning Assessment for Severity of Liver Fibrosis for Chronic HBV Based on Physical Layer With Serum Markers, IEEE Access 7 (2019), pp. 124351-124365

[18] H. Kang, The prevention and handling of the missing data, Korean journal of Anesthesiology, 64 (5) (2013), pp. 402-406.

[19] H.M.K. Ghomrawi, L.A. Mandl, J. Rutledge, M.M. Alexiades, M. Mazumdar, Is there a role for expectation maximization imputation in addressing missing data in research using WOMAC questionnaire? Comparison to the standard mean approach and a tutorial, BMC Musculoskeletal Disorders journal 12 (article no. 109) (2011).

[20] P. Bedossa, T. Poynard, An algorithm for the grading of activity in chronic hepatiti1s C. The METAVIR Cooperative Study Group,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Hepatology 24 (2) (1996), pp. 289–293.
[21] L. Breiman, L. Random Forests, Machine Learning 45 (2001), pp. 5–32.
[22] N. Venkatesan, G. Priya, A Study of Random Forest Algorithm with implementation using Weka, International Journal of Innovative Research in Computer Science and Engineering (IJIRCSE) 1 (issue 6) (2015).
[23] C.Y.J. Peng, K.L. Lee, G.M. Ingersoll, An Introduction to Logistic Regression Analysis and Reporting, The Journal of Educational Research 96 (1) (2002), pp. 3-14.
[24] U. Karthik Kumar, M.B. Sai Nikhil, K. Sumangali, Prediction of Breast Cancer using Voting Classifier Technique, International Conference on Smart Technologies and Application 39 (7) (2012), pp. 6585-6608.
[28] S. Jain. E. Kotsampasakou. G.F. Ecker, Comparing the performance of meta-classifiers a case study on selected imbalanced data sets relevant for prediction of liver toxicity, Journal of Computer Aided Molecular Design 32 (5) (2018), pp. 583-590.
[29] M. Trabelsi, N. Meddouri, M. Meddouri, A New Feature Selection Method for Nominal Classifier based on Formal Concept Analysis, Procedia Computer Science 112 (2017), pp. 186-194.
[30] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The WEKA data mining software: an update, SIGKDD Explorations 11 (1) (2009), pp. 10-18.
[31] University of Waikato: Waikato, New Zeland Weka: Waikato Environment for Knowledge Analysis. http://www.cs.waikato.ac. nz/ml/weka/. Accessed 2 Nov 2010
[32] Mallet, V. Dhalluin-Venier, C. Roussin, M. Bourliere, M.E. Pettinelli, C. Giry, A. Vallet-Pichard, H. Fontaine, S. Pol, The accuracy of the FIB-4 index for the diagnosis of mild fibrosis in chronic hepatitis B, Aliment Pharmacol Ther. 29 (4) (2008), pp. 409-415.
[33] H. Wang, L. Xue, R. Yan, Y. Zhou, M.S. Wang, M.J. Cheng, H.J. Huang, Comparison of FIB-4 and APRI in Chinese HBV-infected patients with persistently normal ALT and mildly elevated ALT, Journal of Viral Hepatitis 20 (4) (2013), e3 – e10.
[34] Z. Yin, J. Zou, Q. Li, L. Chen, Diagnostic value of FIB-4 for liver fibrosis in patients with hepatitis B: a meta-analysis of diagnostic test, Oncotarget 8 (14) (2017), pp. 22944-22953.
[35] N. Papadopoulos, S. Vasileiadi, M. Papavdi, E. Sveroni, P. Antonakaki, E. Dellaporta, E. Koutli, S. Michalea, S. Manolakopoulos, J. Management for Computing, Communication, Controls, Energy and Materials (ICSTM) (2017), pp. 108-114.
[25] S. Kotsiantis, D. Kanellopoulos, P. Pintelas, Handling imbalanced datasets: a review, GESTS International Transactions on Computer Science and Engineering 30 (2006).
[26] A Ali, S.M. Shamsuddin, A.L. Ralescu, Classification with class imbalance problem: a review, Int. J. Advance Soft Compu. Appl. 5 (3) (2013).
[27] López, A. Fernández, J.G. Moreno-Torres, F. Herrera, Analysis of preprocessing vs. cost-sensitive learning for imbalanced classification: open problems on intrinsic data characteristics, Expert System with Koskinas, et al., Liver fibrosis staging with combination of APRI and FIB-4 scoring systems in chronic hepatitis C as an alternative to transient elastography, Annals of Gastroenterology 32 (2019), pp. 498-503.
[36] K. Staufer, E. Halilbasic, W. Spindelboeck, M. Eilenberg, G. Prager, V. Stadlbauer, A. Posch, P. Munda, R. Marculescu, B. Obermayer-Pietsch, et al., Evaluation and comparison of six noninvasive tests for prediction of significant or advanced fibrosis in nonalcoholic fatty liver disease, United European Gastroenterology Journal, 7 (8) (2019), pp. 1113-1123.
[37] R. Fluss, D. Faraggi, B. Reiser, Estimation of the Youden Index and its Associated Cutoff Point, Biometrical Journal 47 (2005), pp. 458-472.