Acıbadem Univ. Sağlık Bilim. Derg. 2021; 12 (3): 625-629
https://doi.org/10.31067/acusaglik.928498

ORIGINAL ARTICLE / ARAŞTIRMA YAZISI Gastroenterology and Hepatology / Gastroenteroloji ve Hepatoloji

Artificial Intelligence to Predict Esophageal Varices in Patients with Cirrhosis

Cem Şimşek[1], İbrahim Emir Tekin[2], Hasan Şahin[2], Taha Koray Şahin[2], Yasemin Hatice Balaban[1]

[1]Hacettepe University Faculty of Medicine, Department of Gastroenterology, Ankara, Turkey
[2]Hacettepe University Faculty of Medicine, Department of Internal Medicine, Ankara, Turkey

Cem ŞİMŞEK, M.D.
Emir TEKİN, M.D.
Hasan ŞAHİN, M.D.
Taha Koray ŞAHİN, M.D.
Yasemin Hatice BALABAN, M.D.

Correspondence: Cem Şimşek
Hacettepe University Faculty of Medicine, Department of Gastroenterology, Ankara, Turkey
Phone: +905342964684
E-mail: cemgsimsek@gmail.com

Received : 27 April 2021
Accepted : 06 June 2021

ABSTRACT

Background: Screening for varices remains as the best strategy to decrease associated mortality that reaches 25%. Diagnostic endoscopy is gold standard but invasive for routine screening. Non-invasive stiffness measurements with elastography is costly and impractical. Non-elastogarphic tests that use available laboratory and clinical variables are feasible but their performance remains inferior to elastography. Non-invasive, accessible and accurate test is needed. Machine learning methods can be used in this sense to provide better diagnostic performances. We aimed to test the ability of a machine learning model to predict esophageal varices in patients with cirrhosis.

Materials and methods: We retrospectively evaluated patients with cirrhosis at the time of their screening upper endoscopies from our institutional database. Demographic, clinical, radiologic, endoscopic and laboratory data was collected. Child-Pugh, APRI, FIB-4, AAR, PCSD tests were calculated for each patient. Gradient boosted machine learning algorithm was constructed for the problem. A logistic regression as well as tests' and model's performances with areas under ROCs were compared to detect presence of esophageal varices.

Results: Study population consisted of 201 patients whom 105 had esopheageal varices which 33 were higher risk. Patients with varices were older, advanced Child stages, larger splenic diameters and higher MELD-Na scores. Composite scores' were as follows: FIB-4 0.57 (0.49-0.65), APRI 0.47 (0.38-0.55), PCSD 0.511 (0.42-0.59), AAR 0.481 (0.39-0.56). Machine learning model's mean AUC to predict varices was 0.68(0.060), F1- score was 0.7 and accuracy was 63%.

Conclusions: Machine learning model outperformed non-invasive tests to predict esophageal varices in cirrhotic patients.

Keywords: esophageal varices, artificial intelligence, machine learning, screening, prediction

Sirozlu Hastalarda Yapay Zeka ile Özofagus Varis Tahmini

ÖZET

Giriş ve amaç: Sirozlu hastalarda özofagus varis taraması, ilişkili mortaliteyi %25'e varan oranlarda azaltmak için en iyi strateji olmaya devam etmektedir. Tanısal üst endoskopi altın standarttır ancak invaziv olması rutin taramayı güçleştirmektedir. Elastografi ile non-invaziv fibrosis ölçümleri maliyetli ve pratik değildir. Mevcut laboratuvar ve klinik değişkenleri kullanan testlerin ise performansları elastografiden daha düşük kalmaktadır. Non-invaziv, erişilebilir ve doğru testler gereklidir. Bu bağlamda varis riskini belirlemek için makine öğrenmesi yöntemleri kullanılabilir. Bu çalışmada, bir makine öğrenme modelinin sirozlu hastalarda özofagus varislerini tahmin etme performansını ve kullanılabilirliğini test etmeyi amaçladık.

Gereç ve yöntem: Kliniğimizin veri tabanından üst endoskopi ile varis taraması yapılan sirozlu hastaları geriye dönük olarak değerlendirdik. Demografik, klinik, radyolojik, endoskopik ve laboratuvar verileri toplandı. Her hasta için Child-Pugh, APRI, FIB-4, AAR, PCSD testleri hesaplandı. Problem için gradyan destekli makine öğrenme algoritması oluşturulmuştur. Özofagus varislerinin varlığını tespit etmek için lojistik regresyon ile testlerin ve modelin ROC'lerin altındaki alanlarla olan performansları karşılaştırıldı.

Bulgular: Çalışma popülasyonu, 105'i özofagus varisi olan ve 33'ü daha yüksek riskli olan 201 hastadan oluşturuldu. Varisli hastalar daha yaşlı, ileri Child evreleri, daha büyük dalak boyutları ve daha yüksek MELD-Na skorlarına sahipti. Testlerin varis olan hastaları tahmin performanslarının AUC değerleri: FIB-4 0,57 (0,49-0,65), APRI 0,47 (0,38-0,55), PCSD 0,511 (0,42-0,59), AAR 0,481 (0,39-0,56) şeklindeydi. Makine öğrenimi modelinin varisleri tahmin etmek için ortalama AUC değeri 0.68(0.060), F1- skoru 0.7 ve doğruluk %63 idi.

Sonuçlar: Makine öğrenimi modellerinin, sirotik hastalarda özofagus varislerini tahmin etmekteki performansı, invazif olmayan testlerle karşılaştırılabilir düzeydeydi.

Anahtar Kelimeler: Karaciğer hastalığı, siroz, özofagus varisleri, yapay zeka, makine öğrenmesi

Variceal bleeding is a major cause of morbidity and mortality in cirrhotic patient. Early identification of varices and primary prophylaxis remains as the most feasible strategy. The gold standard for detecting varices is upper endoscopy but its use is not convenient for repeated screening procedures. Liver stiffness measurements with transient elastography reached performances enough to be implemented in clinical practice as expanded Baveno VI criteria but requires expensive devices along with an experienced operator, thus not also an optimal screening strategy (1). Tests without an elastographic measurements have been proposed but their performance is inferior to elastography. Therefore, a non-invasive but practical test is required to stratify patients for endoscopic screening.

Artificial intelligence is a general term includes several domains of advanced computer programs that can achieve human like cognitive abilities. Machine-learning is a subdomain of artificial intelligence that learns from the data and the problem without needing to be programmed so. These approaches are increasingly being used in virtually every field of medicine as well as hepatology to tackle long-standing problems with their inherent abilities to and integrate a bigger dimensions and extent of data into their solution.

With the need of a screening tool for varices and the promise of machine learning approach, we aimed to test a machine learning model's performance to predict the presence of esophageal varices in patients with cirrhosis. We hypothesize that machine learning's performance will not be inferior to already existing non-invasice clinical/laboratory depedentent scores.

Materials and Methods

Study Design and Patient population

We retrospectively evaluated our endoscopy database for patients who have undergone upper endoscopy for treatment or screening of esophagogastric varices between January 2015 and January 2021. We included patients with an administrative code for cirrhosis or chronic liver diseases (ICD-10, 10th revision of the International Statistical Classification of Diseases and Related Health Problems) who undergone upper endoscopy for the purpose of screening or prophylactic therapy of varices. We confirmed diagnosis of cirrhosis through evaluation of patient charts and radiologic studies. We excluded cases with inaccessible endoscopic, clinical, or laboratory data. Patients with incomplete vital signs were not excluded.

Data Collection and Variables

After confirmation of final patient list, we retrospectively collected data from endoscopy reports, physician notes during inpatient and outpatient encounters, laboratory results and abdominal radiology reports. Patient demographics, vital signs during encounter (Temperature, blood pressure, heart rate, respiratory rate) etiology of liver disease, presence of ascites or hepatic encephalopathy, splenic length in abdominal imaging studies and laboratory values (complete blood count, routine biochemistry, coagulation tests) were collected at the nearest time to upper endoscopy. Child-Pugh scores, Child Classes and MELD-Na scores were calculated. Endoscopy reports were evaluated for the presence of esophageal or gastric varices. If present, esophageal varices were classified as higher- and lower-risk (2)

Machine Learning Models, Feature Selection and Model Training

Adopted machine learning method -Light Gradient Boosting Machine- is an ensemble of multiple decision trees algorithms that learns from each tree to generate a final accurate model of its own (Ke et al. 2017; Chen and Guestrin 2016). We used our database both to train and test the algorithms prediction performance. To increasing the generalizability of our results, we used multiple different splits for training and testing the algorithm. We shuffled the data before every iteration and split it into different training and test sets with four to one ratio that was repeated 50 times. As our population size is limited, we were not able to integrate all variables into the final model that would cause overfitting. We used two feature importance techniques - permutation feature importance and leave-one-out feature importance - to determine which variables to include. Those parameters are selected intuitively rather than using a black box optimizer which can induce overfitting. As the output, mean of 50 models' area under the rule operator curves (AUC) is presented

and standard deviation of the scores is the confidence interval

Outcomes and Statistical Analyses

The characteristics of patient populations was presented with descriptive statistics using median with range for non-parametric continuous variables, mean with standard deviation for parametric continuous variables and ratios with percentages for categorical variables. Patients with and without varices were compared using Mann Whitney U and Chi-square tests when appropriate. A binary logistic regression model was used to find variables that predicts presence of varices. Areas under the ROCs of MELD-Na (3), CTP (4), AST to Platelet Ratio Index (APRI)(5), and Platelet Count to Spleen Diameter (PC/SD)(6), FIB-4(7) scores and AST to ALT ratio were compared to machine learning models for prediction of cirrhosis.

Results

Patient Population

We included 201 patients of clinically or radiologically confirmed cirrhosis. Mean age of the population was 58.0 (16.3). Etiologies of cirrhosis were chronic Hepatitis B, chronic Hepatitis C, non-alcoholic steatohepatitis, alcoholic liver disease, autoimmune liver diseases, Wilson's disease, primary and secondary hemochromatosis, congenital liver diseases, Budd-Chiari syndrome, congenital or acquired hypercoagulatory disorders. Median Child-Pugh score of population was 7 (5-13), 86 cases were Class A, 81 cases were Class B and 34 cases were Class C. Median MELD-Na score of the population was 10 (6-40); 93 patients scores were between 6 and 9, 53 patients scores were between 10 and 19, 23 patients scores were between 20 and 29, and 21 patients scores were equal to or greater than 30 (Table 1).

Varices and Predicting Variables

One-hundred and five patients had esophageal varices as opposed to 96 patients. Of 105 varices, 63 were low-risk and 33 were higher risk. Patients with varices were older (63 vs 54), higher Child-Pugh scores, larger splenic diameters (15.1 vs 13.9) advanced Child stages (64 Child B-C vs. 52 Child B-C)as well as higher MELD-Na scores (19 vs 13).

A binomial logistic regression was performed to ascertain the effects of age, splenic vein diameter, platelet counts and MELD-Na scores on the likelihood that patients have varices. The logistic regression model was statistically significant, $\chi 2(4) = 19.20$, $p < .001$. The model explained 14.0% of the variance in presence of esophageal varices and correctly classified 52.% of cases. Composite scores were calculated and their AUCs to classify presence of varices were as follows: FIB-4 0.57 (0.49-0.65), APRI 0.47 (0.38-0.55), PCSD 0.511 (0.42-0.59), AAR 0.481 (0.39-0.56) (Figure 1).

Table 1. Characteristics of our study population

		Mean (SD) / Number (%)
Age		58 (1)
Gender	Male	97 (48.3%)
	Female	104 (51.7%)
Etiology of Liver Disease	Chronic Hepatitis B	28 (14.4%)
	Chronic Hepatitis C	6 (3.1%)
	Non-alcoholic steatohepatitis	27 (13.9%)
	Alcoholic liver disease	12 (6.2%)
	Cryptogenic	58 (29.9%)
	Autoimmune liver diseases	6 (3.1%)
	Vascular and hypercoagulability	39 (20.1%)
	Malignancy	12 (6.2%)
	Congenital liver diseases	6 (3.1%)
Child Class	Class A	86 (42.8%)
	Class B	81 (40.3%)
	Class C	34 (16.9%)
MELD-Na Group	<10	93 (48.9%)
	19-Oct	53 (27.9%)
	20-29	23 (12.1%)
	>30	21 (11.1%)
Hemoglobin (g/dL)		12.4 (3)
Platelet Count (^3 / mL)		137 (6)
Sodium (mg/dL)		136 (0)
Creatinine (mg/dL)		0.88 (0.04)
ALT (IU/mL)		35 (2)
AST (IU/mL)		56 (5)
ALP (IU/mL)		149 (9)
GGT (IU/mL)		123 (11)
Bilirubin (mg/dL)		686 (392)
INR		2817 (951)

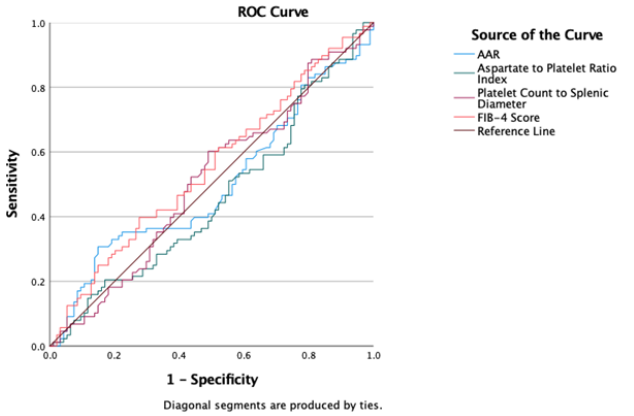

Figure 1. Area under the rule operator curves for APRI, AAR, FIB-4 scores and PCSD ratio to classify patients with cirrhosis.

Model Outputs

Machine learning model's classification performance was tested with prediction of esophageal varices in patients with cirrhosis. Feature selection as described choose following variables: Gender, presence of ascites, presence of encephalopathy, Child-Pugh Score, Platelet counts. Machine learning models mean AUC to predict varices was 0.68(0.060), F1- score was 0.7 and accuracy was 63%. (Figures 2).

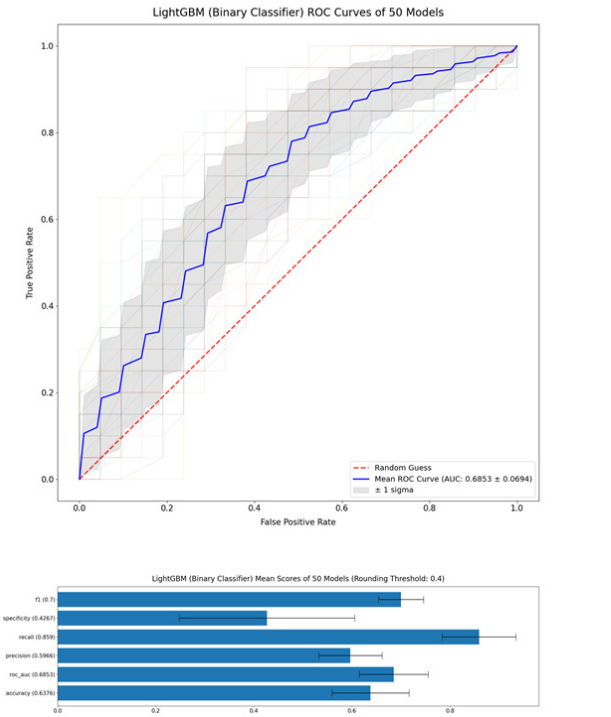

Figure 2. Mean of 50 machine learning models area under rule operator curves to classify patients with cirrhosis.

Table 2. Features of patients with and without cirrhosis

		Varices at Upper Endoscopy		p
		No	Yes	
		Mean (SD)	Mean (SD)	
Age		54 (17)	63 (14)	0.000
Platelet Count		144 (94)	129 (87)	0.235
Splenic largest diameter (cm)		15.1 (3.9)	13.8 (3.5)	0.06
Child Class	Class A	53	33	0.06
	Class B	38	43	
	Class C	14	20	
MELD-Na		13.9252	19.3 (11)	0.006
Aspartate to Platelet Ratio Index		1.94 (4.72)	1.58 (1.84)	0.875
Platelet Count to Splenic Diameter		10.54 (8.86)	10.16 (8.04)	0.893
FIB-4 Score		5.04 (4.08)	5.95 (4.53)	0.65

Discussion

We tested the feasibility of a machine learning model to predict presence of esophageal varices in cirrhotic patients. Our model achieved a higher performance for this task when compared to other composite scores with an AUC of 0.68 which was higher than FIB-4's (0.57), APRI's (0.47), AAR (0.481) and PCSD's (0.511).

Screening for varices is an essential component of clinical management of patients with cirrhosis. Upper endoscopy remaining as the gold standard, current recommendation is the use of non-invasive tests to stratify patients for screening endoscopy. Transient elastography reached sensitivities and specifities over 90% and with the expanded Boveno IV criteria it is now incorporated into clinical practice (1). However, transient elastography is operator dependent and requires costly imaging. In contrast, non-elastographic tests such as APRI score, PCSD ratio, FIB-4 and AAR use readily available laboratory data. However, the tests without elastography have not reached the performance of transient elastography and low to moderate accuracy (8). Previous studies with APRI score

demonstrated specifities between 51%-89% and sensitivities varying between 56% to 71% (9-11). FIB-4 score's and AAR index's performances were similar with sensitivities between 37%-85% and specifities of 64%-81% of (10, 12); sensitivities of 68%-69% and specifities of 34%-89% (8).

Above mentioned non-elastographic scores and indexes use one to three variables to predict a and a complex physiology and a multifactorial condition. Artificial intelligence provides a new perspective to this problem with its ability to integrate greater number and extent of variables to the final decision. As such, there have been several studies using this approach to predict varices. Dong et al created a score using a similar decision tree based machine learning algorithm to create a formula using INR, platelets, BUN, Hemoglobin and ascites. This composite score classified patients with varices with AUC of 0.81 in validation cohort (13).

We acknowledge our studies limitations inherent to retrospective design, small population size, and the use of machine learning methods. Artificial intelligence own specific limitations such as over-fitting regardless of multiple training and test splits as mentioned. Further validation of our model in different and larger datasets is required. We also acknowledge neither AUC of 0.68 of our algorithm nor the sample size of our study is enough to implement artificial intelligence alone by a mean of varix screening but only as a proof of concept for this clinical problem. Moreover, we need to test different algorithms for prediction of varices in different contexts as their pathophysiology, therefore predictive factors, will be presumably different(14).

Knowledge gaps in the management of liver diseases can be targeted with artificial intelligence methods as we already own the required big multimodal data that include radiology, genomics, clinical and laboratory variables. Despite this promise, the future of artificial intelligence in hepatology depends on further efforts and prospective studies.

References

1. Augustin S, Pons M, Maurice JB, Bureau C, Stefanescu H, Ney M, et al. Expanding the Baveno VI criteria for the screening of varices in patients with compensated advanced chronic liver disease. Hepatology. 2017;66(6):1980-8.
2. Garcia-Tsao G, Abraldes JG, Berzigotti A, Bosch J. Portal Hypertensive Bleeding in Cirrhosis: Risk Stratification, Diagnosis, and Management: 2016 Practice Guidance by the American Association for the Study of Liver Diseases (vol 65, pg 310, 2017). Hepatology. 2017;66(1):304-5.
3. Kamath PS, Wiesner RH, Malinchoc M, Kremers W, Therneau TM, Kosberg CL, et al. A model to predict survival in patients with end-stage liver disease. Hepatology. 2001;33(2):464-70.
4. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. Br J Surg. 1973;60(8):646-9.
5. Lin ZH, Xin YN, Dong QJ, Wang Q, Jiang XJ, Zhan SH, et al. Performance of the aspartate aminotransferase-to-platelet ratio index for the staging of hepatitis C-related fibrosis: an updated meta-analysis. Hepatology. 2011;53(3):726-36.
6. Giannini EG, Botta F, Borro P, Dulbecco P, Testa E, Mansi C, et al. Application of the platelet count/spleen diameter ratio to rule out the presence of oesophageal varices in patients with cirrhosis: a validation study based on follow-up. Dig Liver Dis. 2005;37(10):779-85.
7. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology. 2006;43(6):1317-25.
8. Deng H, Qi XS, Guo XZ. Diagnostic Accuracy of APRI, AAR, FIB-4, FI, King, Lok, Forns, and FibroIndex Scores in Predicting the Presence of Esophageal Varices in Liver Cirrhosis A Systematic Review and Meta-Analysis. Medicine. 2015;94(42).
9. Tafarel JR, Tolentino LHL, Correa LM, Bonilha DR, Piauilino P, Martins FP, et al. Prediction of esophageal varices in hepatic cirrhosis by noninvasive markers. Eur J Gastroen Hepat. 2011;23(9):754-8.
10. Stefanescu H, Grigorescu M, Lupsor M, Maniu A, Crisan D, Procopet B, et al. A New and Simple Algorithm for the Noninvasive Assessment of Esophageal Varices in Cirrhotic Patients Using Serum Fibrosis Markers and Transient Elastography. J Gastrointest Liver. 2011;20(1):57-64.
11. Wang JH, Chuah SK, Lu SN, Hung CH, Chen CH, Kee KM, et al. Transient elastography and simple blood markers in the diagnosis of esophageal varices for compensated patients with hepatitis B virus-related cirrhosis. J Gastroen Hepatol. 2012;27(7):1213-8.
12. B CL-GN, De Vinatea-Serrano L, Piscoya A, Segura ER. [Performance of the FIB-4 index in esophageal varices screening in patients with the diagnosis of liver cirrhosis]. Rev Gastroenterol Peru. 2020;40(1):29-35.
13. Dong TS, Kalani A, Aby ES, Le L, Luu K, Hauer M, et al. Machine Learning-based Development and Validation of a Scoring System for Screening High-Risk Esophageal Varices. Clinical Gastroenterology and Hepatology. 2019;17(9):1894-901.e1.
14. Beyazit Y, Ibis M, Purnak T, Turhan T, Kekilli M, Kurt M, et al. Elevated levels of circulating angiotensin converting enzyme in patients with hepatoportal sclerosis. Digestive diseases and sciences. 2011;56(7):2160-5.