Development of a Web Application based on Machine Learning for screening esophageal varices in cirrhosis

Développement d'une Application Web basée sur la Machine Learning pour l'aide au dépistage des varices œsophagiennes au cours de la cirrhose

Soumaya Mrabet[1,3], Kamel Aloui[2], Elhem Ben Jazia[1,3]

1. Service de Gastroentérologie. Hôpital Farhat Hached, Sousse, Tunisie
2. Institut Supérieur d'Informatique et des Techniques de Communication (ISITCom), Sousse, Tunisie
3. Université de Sousse, Faculté de Médecine de Sousse. Tunisie

Abstract

Introduction: Esophageal varices (EV) are a common manifestation of portal hypertension in cirrhotic patients. Upper gastrointestinal endoscopy (UGE) is the gold standard for diagnosing EV. However, it is an invasive examination with a relatively high cost.
Aim: To develop a machine learning model for the prediction of EV in cirrhotic patients.
Methods: This is a cross-sectional observational study including all cirrhotic patients, for whom an UGE was performed, between January 2010 and December 2019. We adopted a structured methodical approach with reference to CRISP-DM (Cross-Industry Standard Process for Data Mining). The different steps carried out were: data collection and preparation, modelization, and deployment of the predictive models in a web application.
Results: We included 166 patients, 92 women (55.4%) and 74 men (44.6%). The mean age was 57.2 years. In UGE, 16 patients (9.6%) did not have EV. Other patients had EV grade 1 in 41 cases (24.7%), grade 2 in 81 cases (24.7%) and grade 3 in 28 cases (16.9%). After the selection phase, among the 36 initial variables, 19 were retained. Three machine learning models have been developed with a performance of 90%.
Conclusions: We developed a machine learning model combining several clinical and para-clinical variables for the predcition of EV in cirrhotic patients.

Key words: artificial intelligence; machine learning; esophageal varices; cirrhosis

Résumé

Introduction: Les varices œsophagiennes (VO) constituent une manifestation fréquente de l'hypertension portale chez les patients cirrhotiques. La fibroscopie digestive haute (FDH) constitue l'examen de référence pour le diagnostic de VO. Cependant, il s'agit d'un examen invasif associé à un coût relativement élevé.
Objectif: Développer une application web basée sur le machine learning et le data mining pour le dépistage de VO au cours de la cirrhose.
Méthodes: Il s'agit d'une étude observationnelle analytique. Nous avons colligé les dossiers des patients suivis pour cirrhose et ayant bénéficié d'une FDH entre Janvier 2010 et Décembre 2019. On a adopté une approche méthodique structurée en se référant au CRISP-DM (Cross-Industry Standard Process for Data Mining). Les différentes étapes réalisées étaient : la collecte et la préparation des données, la modélisation, l'évaluation et le déploiement du modèle prédictif dans une application Web.
Résultats: Nous avons inclus 166 patients, 92 femmes (55,4%) et 74 hommes (44,6%). La moyenne d'âge était de 57,2 ans. A la FDH, 16 patients (9,6%) n'avaient pas de VO. Les autres patients avaient des VO grade 1 dans 41 cas (24,7%), VO grade 2 dans 81 cas (24,7%) et VO grade 3 dans 28 cas (16,9%). Après la phase de sélection, parmi les 36 variables initiales, 19 ont été retenues. Trois modèles de machines learning ont été élaborés :un modèle qui permet de prédire l'absence (1) ou la présence (0) des VO, un modèle qui permet de prédire l'absence de larges VO (0 ou grade 1) ou la présence de larges VO (grade 2 ou 3) et un modèle qui permet de prédire le grade 2 ou 3 des VO. Les performances de ces modèles étaient excellentes avec une précision de 90%. Un site web comportant 3 pages a été crée : une page d'accueil, une page pour l'exploration des données et une page pour le dépistage des VO. L'utilisateur aura le choix entre deux techniques de machine Learning pour la prédiction et la caractérisation des VO : les réseaux de neurones et la régression logistique.
Conclusion: Nous avons élaboré un outil informatique combinant plusieurs variables cliniques et para-cliniques, pour une approche non endoscopique du diagnostic des VO en Tunisie. Des études plus larges seraient nécessaires afin de valider les résultats de notre étude.

Mots clés: Intelligence artificielle, machine learning, varices œsophagiennes, cirrhose

Correspondance
Soumaya Mrabet
Service de Gastroentérologie. Hôpital Farhat Hached, Sousse, Tunisie
Email: mrabetsoumaya99@yahoo.fr

INTRODUCTION

Esophageal varices (EV) are a common manifestation of portal hypertension in patients with cirrhosis. The appearance of EV constitutes an evolutionary turning point in cirrhosis and requires special measures in the management of cirrhotic patients [1]. Despite the progress made in the early diagnosis of EV and the treatment of gastro-intestinal bleeding caused by EV, the mortality rate associated with this complication remains high. Currently it is recommended to screen EV by an upper gastro-intestinal endoscopy (UGE) at the time of diagnosis of cirrhosis and every one to three years after, depending on the grade of varices and the severity of the cirrhosis [2]. UGE is the gold standard for the diagnosis of EV. However, this invasive examination is often poorly tolerated by patients because of the discomfort and perceived risk. In addition, the cost is relatively high [3]. Therefore, the use of non-invasive biological and radiological markers to predict the presence of EV in cirrhosis is necessary [4].

Artificial Intelligence (AI) is defined as «the theories and techniques implemented in order to obtain machines capable of simulating intelligence.» These machines are able to reason, process large amounts of data, discern patterns undetectable by the human eye, understand and analyze patterns, interact with humans, learn progressively in order to continuously improve their performance [5]. The use of artificial intelligence for medicine has recently gained a lot of attention due to advancements in machine learning techniques involving multiple layers of artificial neural networks (ANN) trained on big data [6].

In gastroenterology, AI has been used primarily in digestive endoscopy: characterization of polyps and predicting signs of malignancy. In hepatology, it has been used primarily to assess hepatic fibrosis [7].

The aim of our study is to develop a web application based on innovative approaches of AI for screening EV in cirrhosis. This application is a cognitive computing solution to select patients at high risk of developing EV. Therefore, based on a number of clinical and paraclinical parameters, the patient is classified into one of the following categories: absence of EV, EV grade 1, EV grade 2 or 3. This application allow to professionals to select cirrhotic patients with a high risk of developing EV, avoid unnecessary endoscopic investigations for patients with a low risk of developing EV, improve the quality of care and referral for patients and reduce the human, logistical and financial cost of cirrhotic patients care.

METHODS

This is a cross-sectional observational study, including cirrhotic patients between January 2010 and December 2019, for whom, an UGE was performed.

Study population:

All the patients followed for cirrhosis were included. The diagnosis of cirrhosis was made on clinical, biological, morphological and endoscopic criteria.

We excluded in our study patients on prophylactic treatment for EV disruption, patients with diuretic treatment, patients with end-stage renal disease and patients with specific cause of splenomegaly.

Methodic approach:

In our study, we have adopted a structured methodic approach with reference to CRISP-DM [8] (Cross-Industry Standard Process for Data Mining). This approach contains four steps: Data collection and understanding, Data preparation, modelization and deployment.

• *Data collection and understanding:*

Data collection was carried out from patients' medical records. Four classes have been defined: Absence of EV, EV grade 1, EV grade 2 or 3. The Large EV class means patients with EV grade 2 or 3. Thirty six variables were collected: age, gender, history of diabetes, arterial hypertension, heart disease, hypothyroidism, the presence of ascites, collateral venous circulation, hepatic encephalopathy, the presence of digestive hemorrhage, hepatocellular carcinoma at the the time of diagnosis, the etiology of cirrhosis (chronic hepatitis B, chronic hepatitis C, autoimmune cirrhosis, primary biliary cholangitis, metabolic cirrhosis, vascular cirrhosis, alcoholic cirrhosis or cirrhosis of unknown etiology), biological data (Platelet count, white blood cells, hemoglobin, natremia, ALAT, ASAT, TB, APL, GGT, Prothrombin rate, INR, albuminemia, creatinine) and mophological data assessed by abdominal ultrasound (diameters of the spleen and of the portal vein). We used also the CHILD score and MELD score to evaluate the severity of cirrhosis.

• *Data preparation:*

This step involved preprocessing and cleaning the raw data into a tidy dataset by exploring it, handling missing data, and removing outliers (noise). Missing data was replaced with the most frequent value. Redundant data has been deleted. Then, the relevant data was selected for the development of the predictive models.

• *Modelization:*

During this step, the development of predictive models is based on machine learning algorithms allowing the series of biomedical measurements relating to a patient to be associated with one of the following classes: Absence of EV, EV grade 1, EV grade 2 and EV grade 3. We used two models: logistic regression and artificial neural network (ANN). The performances of each model were evaluated using sensitivity, specificity and precision. The most efficient models were selected and validated on 10 patients.

• *Deployment*

During this last step, the predictive models were deployed in a Web application using the Streamlit application. This application was installed using the ANACONDA browser for windows by following the steps described on the following link: How to clean install Streamlit — Streamlit 0.70.0 documentation

RESULTS

Patient characteristics

166 patients were included: 92 women (55.4%) and 74 men (44.6%), with an average age of 57.2 years (between 17 and 88 years).The medical history found in these patients was diabetes in 47 cases, high blood pressure in 32 cases, heart disease in 14 cases and hypothyroidism in 10 cases. The most frequent complication of cirrhosis was ascitic decompensation found in 113 patients (68.1%). Viral B cirrhosis found in 51 cases (30.7%) was the most frequent etiology. Eighty nine patients (53.6%) were classified CHILD B. The mean MELD score was 6.1 points (between 6 and 34 points). These different characteristics are summarized in Table 1.

Table 1. Characteristics of patients

	Frequency (n)	Rate (%)
Mean age (years)	57.2	
Gender:		
Female	92	55.4
Male	74	44.6
Medical history		
Diabetes	47	28.3
High blood pressure	32	19.3
Hypothyroidism	10	6
Complications of cirrhosis		
Ascites	113	68.1
Collateral venous circulation	68	37.8
Hepatocellular carcinoma	33	18.3
Hepatic encephalopathy	18	10
Gastrointestinal bleeding	12	6.7
Etiology of cirrhosis		
Chronic hepatitis B	51	30.7
Chronic hepatitis C	28	16.9
Autoimmune hepatitis B	10	6
Primary biliary cirrhosis	17	9.4
Metabolic cirrhosis	19	10.6
Vascular cirrhosis	5	2.8
Alcoholic cirrhosis	8	4.4
Unknown etiology	30	16.7
Severity of cirrhosis		
CHILD score: A / B / C	39 / 89 / 38	23.5 / 53.6 / 22.9
MELD score (mean value)	6.1 (6-34)	
Abdominal Ultrasound		
Mean spleen diameter (mm)	152	
Mean portal vein diameter (mm)	14.3	
UGE		
Absence of EV	16	9.6
EV grade 1	41	24.7
EV grade 2	81	48.8
EV grade 3	28	16.9

Biological Data:

The various biological data of the patients are shown in Table 2, with the mean, the minimum and maximum values.

Morphological Data:

The average spleen diameter was 152.5mm (between 100 and 230mm). The average diameter of the portal vein was 14.3 mm (between 7 and 22 mm). In UGE, 16 patients (9.6%) did not have EV. The other patients had EV grade 1 in 41 cases (24.7%), EV grade 2 in 81 cases (24.7%) and EV grade 3 in 28 cases (16.9%).

Table 2. Biological Data of patients

	Minimum	Maximum	Average	SD
Platelet count (elements/mm3)	23 000	284 000	109 469.8	63 208.2
WBC (elements/mm3)	800	19000	5587.9	3455.1
Hemoglobin (g/dl)	4	16.5	10.4	2.5
Na+ (mmol/l)	119	146	135.9	4.8
ALAT (UI/L)	5	334	47.4	45.4
ASAT (UI/L)	13	537	77.1	70.1
TB (micromol/l)	6	425	41.9	55.6
APL (UI/L)	0	836	184.5	146.3
GGT (UI/L)	9	865	110.1	140.2
Prothrombin rate (%)	13	100	59.1	18.1
INR	1.00	6.1	1.6	0.6
Albuminemia(g/l)	15.1	45.0	28.1	5.4
Creatinine(micromol/l)	10	703	78.5	73.1

SD : Standard deviation WBC :White blood cells

Elaboration of models and web application:

After the selection phase, among the 36 variables, 19 were retained. These variables are the most relevant for the prediction of EV based in the univariate analysis. The variables selected are shown in Table 3.

Table 3. Relevant variables for the prediction of Esophageal Varices

1-Age
2-High blood pressure
3-Cardiac disease
4-Chronic hepatitis B
5-Gastro-intestinal Bleeding
6-Collateral veinous circulation
7-platelet count
8-Prothrombin rate
9-ASAT
10-ALAT
11-GGT
12-PAL
13-BT
14-White blood cells
15-Creatinin
16-INR
17-MELD Score
18-Spleen diameter
19-Potal vein diameter

Three machine learning models were developed: A model that predicts the absence (1) or presence (0) of EV, a model that predicts the absence of large veins (0 or grade 1) or the presence of Large EV (grade 2 or 3) and a model that predicts grade 2 or 3 EV. Figure 1 summarizes the decision strategy based on the first two models. The user will have the choice between two machine learning techniques for the prediction and characterization of EV: neural networks or Logistic regression. The performance of these models was excellent with an accuracy of 90%.

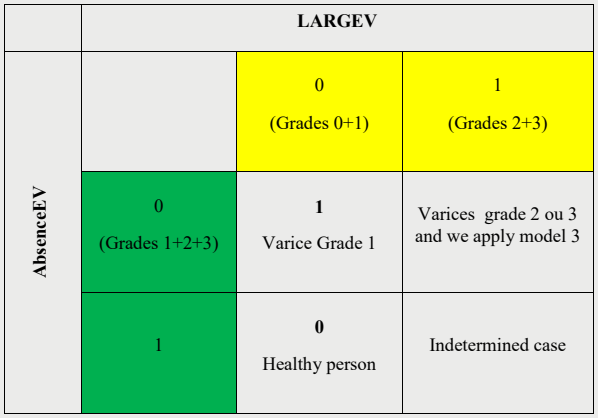

Figure 1. Representation of predictive models

Three machine learning models have been developed:

a. A model that predicts the absence (1) or presence (0) of varicose veins.

b. A model that predicts the absence of large varices (0 or grade 1) or the presence of large VO (grade 2 or 3).

c. A model that predicts grade 2 or 3 EV

We created a website with three pages: A home page that contains the title as well as the presentation of the project. A data exploration page that allows data visualization, with graphical presentations in the form of histograms and curves. A page for the prediction of EV. In this page, the user will have the choice between two learning machine methods (the neural network or logistic regression). The user can also choose the way of entering patient data (graphically, or using a vector of values).

DISCUSSION

We developed three machine learning models combining several clinical and para-clinical variables, for the prediction of EV in cirrhotic patients in order to reduce the use of endoscopy. The use of AI and the integration of different factors in predictive models of EV has only been carried out by a few authors. In the Hong W e al study, which included 197 patients with viral B cirrhosis, the authors developed a non-endoscopic tool for the diagnosis of EV, based on ANN method. In this study, fourteen variables were relevant for the presence of EV. Four variables (platelet count, ascites, diameter of the spleen and diameter of the portal vein) were significantly associated with the presence of EV. The model thus developed had a positive predictive value of 90% and a negative predictive value of 80.8% [9]. In another study, Moulion JR et al developed a non-endoscopic approach to the diagnosis of EV of in a population of sub-Saharan Black Africa, including 98 cirrhotic patients. The method of classification and regression by map made it possible to build a predictive tree for EV. Variables associated with the severity of EV were prothrombin, hemoglobin, albumin, leukocyte, serum creatinine, and spleen size. These factors were used to construct the tree predicting the presence of EV [3]. In a study including 238 patients, Dong T et al, developed an algorithm «the EVendo score» based on the international normalized ratio (INR), aspartate aminotransferase (ASAT) level, platelet count, urea, hemoglobin, and the presence of ascites, using the Random Forest that is a machine learning method for classification with the construction of many decision trees. Each tree consists of nodes made up of variables selected at random. The predictor variable that provides the best allocation is used to form a binary division at a node, and the process is repeated until all predictor variables are used. At the end of this process, a single decision tree is created, and this method is repeated to create several decision treesUsing a composite score of all of these variables, clinicians can quickly and reliably determine which patients are likely to have EV with a sensitivity of approximately 95%. [4].

Stefanescu et al. [10] were the first to consider the technique step by step, using a combination of Lok score (which they had previously validated), LS (liver stiffness) and SS (splenic stiffness). Their algorithm included a first step in which patients with LS <19 kPa and Lok score <0.6 had a low risk of having large EV and those with LS> 38 kPa and Lok score> 0.8 had a high risk to have large EV. In the intermediate zone or the gray zone, splenic elasticity was used as a second step, with a cut off value of 53 kPa. This algorithm provided an accuracy of 78%. The authors concluded that their approach could be used to exclude large EV, but their algorithm provided only a low rate of patients in whom endoscopy can be avoided (24%). In 2017, Pateu E. et al developed another machine-learning model, called VariScreen. A blood test called CirrhoMeter was used initially to identify patients at low or high risk for large varicose veins. Patients belonging to the undetermined area of the CirrhoMeter had endoscopy as a second step. Using VariScreen the rate of patients who avoided endoscopy was 69% with a missed detection rate of large varicose veins of 5.6% [11]. Combinations and multitest algorithms can provide clinically relevant information but still require validation in some cases.

Our study has some limitations. The retrospective character which exposes to the risk of selection and information bias. The relatively small size of our sample reduces the statistical power of our analyzes and makes our results difficult to generalize.

CONCLUSIONS

In our work, we adopted a structured methodical approach by referring to CRISP-DM which contains the following steps: data collection and understanding, data preparation, modeling, assessment and deployment. Of the 36 variables collected, 19 were retained for the development of predictive models. The user will have the choice between two machine-learning techniques for the prediction and characterization of EV: neural networks and logistic regression. The performance of these models was excellent with an accuracy of 90%. During the deployment phase, a website was created using the Streamlit application. The main limitation of our study is the relatively small size of the sample. Thus, this tool deserves to be tested prospectively on a large sample of patients in order to assess its reliability and validity

Prospective work on a larger sample, with more patients who have not yet reached an advanced stage of the disease is needed in order to improve this tool, to verify its diagnostic accuracy and to test its applicability in current practice.

Abbreviations list:

AI: Artificial intelligence

ALAT: Alanine aminotransferase

ANN: Artificial neural networks

APL: Alcaline phosphatase

ASAT: Aspartate aminotranferase

EV: Esophageal varices

GGT: Gamma glutamyl transferase

INR: International normalized ratio

TB: total bilirubin

UGE: Upper gastro-intestinal endoscopy

REFERENCES

1. Sanyal AJ, Bosch J, Blei A, et al. Portal hypertension andits complications. Gastroenterology 2008;134:1715—28.
2. De Franchis R. Revising consensus in portal hypertension:report of the Baveno V consensus workshop on methodologyof diagnosis and therapy in portal hypertension. J Hepatol2010;53:762—8.
3. Moulion Tapouh JR, Njoya O, Monabang Zoé C, el al. Approche non Endoscopique du Diagnostic des Varices OEsophagiennes d'Origine Cirrhotique dans une Population d'Afrique Noire Subsaharienne. Health Sci. Dis 2015;16.
4. Dong T, Kalani A, Aby E, et al. Machine Learning-based Development and Validation of a Scoring System for Screening High-Risk Esophageal Varices. Clinical Gastroenterology and Hepatology 2019;17:1894–1901
5. Microsoft Experiences, Tout savoir sur l'Intelligence Artificielle, available on: https://experiences.microsoft.fr/business/intelligenceartificielle- ia business/comprendre-utiliser-intelligence-artificielle/
6. Seong Ho P, Kyung-Hyun D, Sungwon K, et al. What should medical students know about artificial intelligence in medicine? J Educ Eval Health Prof 2019;16:18.
7. Adadi A , Adadi S ,Berrada M, et al. Gastroenterology Meets Machine Learning. Hindawi Advances in Bioinformatics 2019; 2019
8. Guide CRISP-DM de IBM SPSS Modeler.
9. Hong W, Ji Y, Wang D, et al. Use of artificial neural network to predict esophageal varices in pa¬tients with HBV related cirrhosis. Hepat Mon. 2011;11(7):544.
10. Stefanescu H, Radu C, Procopet B, et al. Noninvasive menage a trois for the predictionof high-risk varices: stepwise algorithm using lok score, liverand spleen stiffness. Liver Int 2015;35:317—25.
11. Pateu E, Oberti F, Cales P, et al. The noninvasive diagnosis of esophageal varices and its application in clinical practice. Clin Res Hepatol Gastroenterol 2018;42(1), 6-16.