Prediction Analysis of Esophageal Variceal Degrees using Data Mining: Is Validated in Clinical Medicine?

By Abd Elrazek Mohammad Aly Abd Elrazek & Hamdy Mahfouz
Al-Azhar University, Egypt

Abstract - The objective of this study is to assess the feasibility of a data mining association analysis technique in early prediction of esophageal varices in cirrhotic patients and prediction of risky groups candidates for urgent interventional procedure. A manuscript titled “Detection of Risky Esophageal varices using 2D U/S: when to perform Endoscopy”, published in The American Journal of The Medical Science on 21Th of December 2012, to our knowledge it was the first prospective study to assess the degree of esophageal varices by 2D ultrasound using the data mining statistical computed analysis in 673 patients. A descriptive model was generated using a decision tree algorithm (Rapid Miner, version 4.6, Berlin, Germany), the over all accuracy was 95%. Following another 59 patients using statistical analysis to determine the association between esophageal variceal degrees detected by Ultrasound in comparable to Upper Endoscopy, was done. Categorical data were compared using the $\chi^2$ test, where as continuous variables were compared using Student’s $t$ test. The comparative results accuracy of both two studies was 97.9%.

GJCST-C Classification : H.2.8
Prediction Analysis of Esophageal Variceal Degrees using Data Mining: Is Validated in Clinical Medicine?

Abd Elrazek Mohammad Aly Abd Elrazek & Hamdy Mahfouz

Abstract - The objective of this study is to assess the feasibility of a data mining association analysis technique in early prediction of esophageal varices in cirrhotic patients and prediction of risky groups candidates for urgent Interventional procedure. A manuscript titled "Detection of Risky Esophageal varices using 2D U/S; when to perform Endoscopy", published in The American Journal of The Medical Science on 21th of December 2012, to our knowledge It was the first prospective study to assess the degree of esophageal varices by 2D ultrasound using the data mining statistical computed analysis in 573 patients. A descriptive model was generated using a decision tree algorithm (Rapid Miner, version 4.6, Berlin, Germany), the over all accuracy was 95%. Following another 59 patients using statistical analysis to determine the association between esophageal variceal degrees detected by Ultrasound in comparable to Upper Endoscopy, was done. Categorical data were compared using the $\chi^2$ test, where as continuous variables were compared using Student's $t$ test. The comparative results accuracy of both two studies was 97.9%.

I. INTRODUCTION

Although the rate of mortality from GI bleeding episodes has decreased with improved endoscopic and radiologic techniques together with new pharmacologic therapies, the 13% to 20% mortality implies the clinical importance, whatever most of mortalities due variceal rupture or leakage [1]. Concern has been raised about the prediction of Esophageal varices for those expected to develop gastrointestinal bleeding in the future, analyzing data by Data mining programs discovered the significant important leading mortality factors; risky Esophageal varices [2]. Thus the major challenge of biomedical data mining over the next 5-10 years is to make these systems useful to biomedical researchers [3]. This paper discuss the utility usefulness of data mining as a predictor statistical tool used to predict esophageal variceal degrees in cirrhotic patients, through measuring the intra-abdominal esophageal wall thickness, by non invasive ultrasound technique.

II. CLINICAL VIEW

a) Causes of Upper GI bleeding

The most common causes of upper gastrointestinal bleeding are: Gastric and/or duodenal ulcers, Esophagogastric varices with or without portal hypertensive gastropathy, Esophagitis, Erosive gastritis/duodenitis, Mallory-Weiss syndrome, Angiodysplasia, Mass lesions (polyps/cancers), Dieulafoy's lesion.[4].

Helicobacter pylori infection, Non Steroidal anti-inflammatory drugs (NSAIDs), physiologic stress, and excess gastric acid are major risk factors for bleeding peptic ulcers. Reduction or elimination of these risk factors reduces ulcer recurrence and rebleeding rates, whatever Esophagogastric varices develop as a consequence of portal hypertension Less common causes of upper gastrointestinal bleeding include Hemobilia, Hemosuccus pancreaticus, Aortoenteric fistula and Cameron lesions.[5,6].

b) Causes of Esophageal Varices

Esophageal varices are swollen blood vessels in the esophagus, which is the tube that connects the mouth to the stomach. Esophageal varices often happen in people with serious liver disease, called "cirrhosis." As a result of portal hypertension [Figure1], Cirrhosis is a medical term, usually used to describe a diseased liver that has been severely scarred injury. There are numerous causes of liver cirrhosis, the two most common causes of cirrhosis worldwide are alcoholic liver disease and viral hepatitis C, which together accounted for approximately one-half of patients on the liver transplantation wait list worldwide whatever portal hypertension can also be present in the absence of cirrhosis, a condition referred to as "noncirrhotic portal hypertension". The causes of non-cirrhotic portal hypertension can be divided into prehepatic, intrahepatic and post hepatic (presinusoidal, sinusoidal and post sinusoidal causes).
c) **Diagnostic Techniques**

Upper gastrointestinal endoscopy remains the gold standard for the diagnosis of esophagastroduodenal lesions, despite its limitations. Early endoscopy (within 24 hours) is recommended for most patients with acute UGI bleeding, though whether early endoscopy affects outcomes and resource utilization is unsettled [7].

d) **Non-Invasive Diagnoses**

The identification of Cirrhotic patients with esophageal varices using non-invasive means has been attempted using different clinical, laboratory and radiological approaches, unfortunately many of these approaches still controversial. Recently non-invasive 2D U/S can detect esophageal variceal degrees in cirrhotic patients with high accuracy, through measuring the intra-abdominal wall thickness of the esophagus, there were proportional relations between esophageal wall thicknesses detected by ultrasound and the esophageal variceal degrees diagnosed by Upper Endoscopy.

III. **Analysis by Data Mining**

The basic classification is based on supervised algorithm, whatever algorithms are applicable for the input data, the accuracy of each algorithm could be changed according to data nature. In clinical medicine, there are different laboratory, clinical and radiological factors determine the progress of each disease. Identification of important and risky factors is of great importance to predict outcome in each disease condition. Thus the major challenge of biomedical data mining is to make these systems useful to biomedical researchers. The decision tree analysis was performed using Intelligent Miner software (Rapid Miner, Berlin, ver.4.6, Germany), which can automatically search a data set to find the optimal classification variables leading to the building of a decision tree algorithm. Briefly, all items derived from the patients were evaluated to determine which variables and cutoff points might produce the most
significant division into two subgroups; group with risky esophageal varices when esophageal wall thickness >6.5 mm and not risky group when esophageal wall thickness ≤ 6.5 mm.

IV. Validation Accuracy

In our previous study the algorithm was selected by evaluating each supervised machine learning algorithms by using supervised learning assessment (10-folds-cross validation) on the training set, we chose the best test applicable to our clinical data, accordingly we used Naïve-base test, and the overall accuracy we obtained was 95%.

We followed 59 patients presented with portal hypertension as a result of End stage-liver cirrhosis; the thicknesses of esophageal walls were measured using 2D U/S. All patients underwent diagnostic Esophagogastroduodenoscopy to estimate corresponding degrees of varices. According to the decision tree algorithm we obtained from the previous study, the esophageal wall thickness > 4.2 mm with inner wall irregularities, should be Esophageal varices with variable degrees. Esophageal wall thickness <4.2mm measured by U/S without inner wall irregularities, indicating no varices, if the esophageal wall thickness > 6.5mm it should be risky esophageal varices (Grade III or IV varices), accordingly upper endoscopy should be performed early as a prophylactic intervention before further bleeding.

V. Results

(out of 59 patients presented with Liver cirrhosis and portal hypertension, only 53 patients had varices when esophageal wall thicknesses were > 4.2 and 6 patients had not varices when esophageal wall thicknesses were <4.2mm, all measured by U/S and diagnosed by upper Endoscopy; (100 % Comparative validation result for detection of Esophageal varices), whatever 48 patients had esophageal wall thicknesses > 6.5mm; risky esophageal varices, only 45 patients had risky esophageal varices diagnosed with Upper endoscopy; (93.75% Comparative validation result for detection of risky Esophageal varices). Categorical data were compared using the $\chi^2$ test, where as continuous variables were compared using Student’s $t$test.

The overall comparative validation accuracy was 97.9 %. Table [1].

<table>
<thead>
<tr>
<th>Esophageal wall Thickness</th>
<th>Cases</th>
<th>Result</th>
<th>Comparative</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 4.2 mm ; No varices</td>
<td>6</td>
<td>6</td>
<td>100%</td>
</tr>
<tr>
<td>&gt;4.2 mm*; Varices</td>
<td>53</td>
<td>53</td>
<td>100%</td>
</tr>
<tr>
<td>&gt;6.5 mm; Risky varices</td>
<td>48</td>
<td>45</td>
<td>93.75%</td>
</tr>
<tr>
<td>Over All Accuracy</td>
<td></td>
<td></td>
<td>97.9%</td>
</tr>
</tbody>
</table>

Table 1 : Describing the Comparative results between Analyses obtained by Data mining Computed Software; Rapid Miner ver.4.6 Berlin, Germany and results obtained from the current study

* Inner esophageal wall irregularities are confirmatory leading signs for esophageal varices with variable degrees.

VI. Conclusion

Based on the available evidence, prediction identification of esophageal variceal degrees in cirrhotic patients presented with manifestations of portal hypertension using non-invasive 2D ultrasound could be a very helpful tool saving time and money. Data mining shed light on the most significant predictors-related esophageal wall thicknesses in 673 patients with very high accuracy; identification high risky groups needed urgent interventional Endoscopy; less complication, better out come and decrease mortality.

Data mining would be the coming statistical evolution in clinical medical data statistical analysis with reasonable limitation.

VII. Limitation of the Study

Given the small sample size of 59 patients (8.7%) in comparable to 673 reported in the pervious study, the validation accuracy might be changed if we apply more patients in the future, whatever our clinical experience played a major role in assessing the information mentioned above, in our point of view our results should be confirmed with more evidence-based criteria using independent laboratory and clinical factors all together.

References Références Referencias


