

October 20-22, 2014 DoubleTree by Hilton Hotel Chicago-North Shore Conference Center, USA

Application of data mining techniques to explore predictors of HCC in Egyptian patients with HCV related chronic liver disease

Dalia A Omran, Abu Bakr A Awad, Mahasen A Mabrouk , Ahmad Fouad and Ashraf Omar Cairo University, Egypt

Background & Aims: Hepatocellular carcinoma (HCC) is the second most common malignancy in Egypt. Data mining is a method of predictive analysis which can explore tremendous volumes of rich information to discover hidden patterns and relationships. We aim to develop a non-invasive algorithm for prediction of HCC. This algorithm should be economical, reliable, easy to apply and acceptable by domain experts.

Methods: This cross sectional study enrolled 315 patients with hepatitis C virus (HCV) related chronic liver disease (CLD); 135 HCC, 116 cirrhotic patients without HCC and 64 patients with chronic hepatitis C. Using data mining analysis, we constructed decision tree learning algorithm to predict HCC.

Results: Decision tree algorithm was able to predict HCC with recall (sensitivity) 83.5% and precession (specificity) 83.3% using only routine data. The correctly classified instances were 259 (82.2%), and the incorrectly classified instances were 56 (17.8%). Out of 29 attributes, serum alpha fetoprotein (AFP), with an optimal cutoff value of \geq 50.3 ng/ml was selected as the best predictor of HCC. To a lesser extent, male sex, presence of cirrhosis, AST \geq 64U/L, and ascites were variables associated with HCC.

Conclusion: Data mining analysis explores data to discover hidden patterns and enables the development of models to predict HCC utilizing routine data as an alternative to CT and liver biopsy. This study has highlighted a new cutoff for AFP (\geq 50.3 ng/ml). Presense of a score of > 2 risk variables (out of 5) can successfully predict HCC with a sensitivity of 96% and specificity of 82%.

daliaomran2007@yahoo.com; daliaomran@kasralainy.edu.eg