



# Models of logistic regression analysis, support vector machine, and back-propagation neural network based on serum tumor markers in colorectal cancer diagnosis

B. Zhang<sup>1</sup>, X.L. Liang<sup>2</sup>, H.Y. Gao<sup>2</sup>, L.S. Ye<sup>2</sup> and Y.G. Wang<sup>1</sup>

<sup>1</sup>Training Department, Third Military Medical University of Chinese PLA, Shapingba, Chongqing, China

<sup>2</sup>Medical Department, Military General Hospital of Beijing PLA, Dongcheng, Beijing, China

Corresponding author: Y.G. Wang  
E-mail: yungui.wang@aol.com

Genet. Mol. Res. 15 (2): gmr.15028643

Received March 23, 2016

Accepted April 11, 2016

Published May 13, 2016

DOI <http://dx.doi.org/10.4238/gmr.15028643>

**ABSTRACT.** We evaluated the application of three machine learning algorithms, including logistic regression, support vector machine and back-propagation neural network, for diagnosing congenital heart disease and colorectal cancer. By inspecting related serum tumor marker levels in colorectal cancer patients and healthy subjects, early diagnosis models for colorectal cancer were built using three machine learning algorithms to assess their corresponding diagnostic values. Except for serum alpha-fetoprotein, the levels of 11 other serum markers of patients in the colorectal cancer group were higher than those in the benign colorectal cancer group ( $P < 0.05$ ). The results of logistic regression analysis indicated that individual detection of serum carcinoembryonic antigens, CA199, CA242, CA125, and CA153 and their combined detection was effective for diagnosing colorectal cancer. Combined detection had a better diagnostic effect with a sensitivity of 94.2% and

specificity of 97.7%; combining serum carcinoembryonic antigens, CA199, CA242, CA125, and CA153, with the support vector machine diagnosis model and back-propagation, a neural network diagnosis model was built with diagnostic accuracies of 82 and 75%, sensitivities of 85 and 80%, and specificities of 80 and 70%, respectively. Colorectal cancer diagnosis models based on the three machine learning algorithms showed high diagnostic value and can help obtain evidence for the early diagnosis of colorectal cancer.

**Key words:** Back-propagation neural network; Colorectal cancer; Logistic regression analysis; Serum marker; Support vector machine