Application of logistic regression in combination with multiple diagnostic tests for auxiliary diagnosis of nasopharyngeal carcinoma

Shi-Qiang Jiang1 and Qing Liu2,3,*

1Department of Medical Statistics and Epidemiology, School of Public Health, Sun Yat-sen University, Guangzhou, Guangdong P.R. China; 2Department of Epidemiology; Cancer Center; Sun Yat-sen University; Guangzhou, Guangdong P.R. China

Key words: nasopharyngeal carcinoma (NPC), Epstein Barr virus (EBV), diagnostic parameters, logistic regression, combination test, sensitivity, specificity

Background and Objective: Although there are many markers for the clinical diagnosis of nasopharyngeal carcinoma (NPC), the efficacy of most of the markers for the early diagnosis is poor. This study was to evaluate the diagnostic value of VCA/IgA, EA/IgA, EBV DNA, EBNA1/IgA, EBNA1/IgG and ZTA/IgG for NPC, as well as to screen out an optimized combination using the logistic regression to increase diagnostic accuracy of NPC. Methods: Eight-one newly pathologically diagnosed NPC patients prior to treatment and 89 health cases from routine physical checkups were entered into the study. Epstein-Barr virus (EBV) DNA was detected by quantitative real-time PCR; VCA/IgA and EA/IgA were assessed by immunofluorescence assays (IFA). The receiver operating characteristic (ROC) curve and the area under the curve were used to evaluate the diagnostic value of a single test or combined tests for NPC, thus to decide the cut-off value. The logistic regression model was used to combine the results from multiple tests to increase diagnostic accuracy. Results: Comparing to the routine parallel sequential test, the logistic regression in combination with multiple diagnostic tests achieved higher diagnostic specificity and sensitivity for NPC.

Patients and Methods

Patients. Peripheral blood samples were collected from 81 newly diagnosed NPC patients in the outpatient clinic of Sun Yat-sen University Cancer Center from August 2006 to December 2006. There were 59 males and 22 females, aged from 16 to 74 years, with a median age of 46 years. All patients underwent nasopharyngoscopy, and were pathologically confirmed by excisional biopsy, which is regarded as the "gold standard" for NPC diagnosis. Eighty-nine patients who lived in Sihui City, Guangdong Province and underwent health checkups were selected as control, among which 41 were males and 48 were females, aged from 26 to 64 years, with a median age of 40 years.

Measurement parameters and methods. Epstein-Barr (EB) viral DNA was measured using real-time fluorescent quantitative PCR, and the copy number of EB virus (EBV) was calculated.
EBV VCA/IgA and EA/IgA were assessed using indirect immunoenzyme staining. EBNA1/IgG, EBNA1/IgA as well as Zta/IgG (serum dilutions: 1:1000, 1:100 and 1:100, respectively) were measured using enzyme linked immunosorbent assay (ELISA). Diagnostic cut-off values of the six markers were as follows respectively: VCA/IgA (1:10), EA/IgA(1:10), EBV DNA (2000 copies/ml), EBNA1/IgA(1), EBNA1/IgG (1), Zta/IgG (1). Due to many tests were performed in this study, some samples were not sufficient to complete all testing items. EBNA1/IgG, EBNA1/IgA and ZTA/IgG assays were performed in only 49 samples.

Statistical analyses. Data were statistically analyzed using SAS9.1 software package. An ROC curve was plotted using PROC LOGISTIC and PROC GPLOT ROC programs of the SAS statistical software package. Area under the ROC curve was calculated using a nonparametric method. Comparison of area under the ROC curve was performed using the %ROC macro.

Results

Diagnostic value of each marker. VCA/IgA (95.06%), EBV DNA (83.95%) and EBNA1/IgA (81.63%) showed relatively higher sensitivity among the six parameters, whereas EA/IgA (94.38%), EBNA1/IgG (95.51%) and Zta/IgG (89.89%) showed higher specificity for the diagnosis of NPC (Table 1).

Diagnosis outcomes of combination of two markers tested by two diagnostic methods. The diagnosis accuracy of a combination of two markers is shown in Table 2. “AND” denotes serial testing, which means that only a positive result on both tests would be considered positive; “OR” denotes parallel testing, which means that only negative results on both tests would be considered negative.

Discussion

Neither parallel testing nor serial testing could improve the sensitivity and specificity for the diagnosis of NPC at the same time. The combination test can incorporate the information detected by multiple markers using the multivariate statistical model for diagnosis evaluation. These statistical models include the multiple linear regression model, the linear discriminant model, the logistic regression model and so on. The multiple linear regression and the linear discriminant models have relatively stringent requirements on the
distribution characteristics of data, whereas the logistic regression is a relatively ideal statistical model used for problem assessment in multi-index tests. As the statistical relationship between the variable and the response variable is fixed in the logistic regression model, once the adoption value of a group of independent variables is given, the type of the individual probability can be identified; or in other words, the interpreted result is not correlated to the distribution type of the independent variable. This feature makes the logistic regression model more stable than conventional discriminant analysis, and makes it a very good system to handle covariates. In this study, we used the logistic regression to set up a model for the combination diagnosis of two NPC markers, compared the obtained results with those of parallel testing and serial testing, and found that the sensitivity and specificity were both improved using the logistic regression.

Previous studies revealed that the combination of EBV antibody VCA/IgA with EA/IgA yielded a relatively high diagnostic accuracy. Therefore, these two markers have been regarded as essential reference markers for clinical diagnosis of NPC and for screening of high-risk population for many years. In recent years, many researchers have adopted single detection modality or the comparison modality with VCA/IgA and assessed some new markers for EBV detection, including EBNA1/IgA, EBNA1/IgG, ZTA/IgG and free EBV DNA copy number. However, there are few reports regarding to comprehensive comparisons among these markers, as well as the predictive value of combination diagnosis of these markers.

We found that the sensitivity of VCA/IgA was relatively high, which is similar to other studies. Additionally, the specificity of EA/IgA and EBNA1/IgG were also high. Due to the high cost and long duration required for the quantitative measurement of EBV DNA, it is not suitable for outpatient clinic use. We only assessed the diagnostic value of the combination of two markers in our study, because detecting excessive markers would increase the financial burden of outpatients. Two groups of combinations were found to achieve high predictive values for the diagnosis: EBV DNA + EBNA1/IgA and EBNA1/IgA + VCA/IgA. As the assay of EBV DNA is complicated and costly, whereas the measurements for EBNA1/IgA and VCA/IgA are rather simple and inexpensive, the combination of EBNA1/IgA + VCA/IgA is recommended for early diagnosis of NPC.

Adopting the logistic regression to incorporate information measured by multi-factors, and determining the diagnostic cut-off value according to the ROC curve could improve the diagnostic sensitivity and specificity for NPC at the same time. This method can also be applied for information analysis and diagnosis of other diseases.

Acknowledgements

We would like to thank the Department of Nasopharyngeal Carcinoma, Sun Yat-sen University Cancer Center for helping collect the cases and also thank the Cancer Institute of Sihui City for providing us healthy controls.

Grant: National “11th Five-Year” Key Technology R&D Program (No. 2006BAI02A11).

References

Application of logistic regression in combination with multiple diagnostic tests for auxiliary diagnosis of nasopharyngeal carcinoma


