Comparison of Fuzzy Inference, Logistic Regression, and Classification Trees (CART)

Prediction of Cervical Lymph Node Metastasis in Carcinoma of the Tongue

G. Schwarzer¹, T. Nagata², D. Mattern³, R. Schmelzeisen⁴, M. Schumacher¹

¹Institute of Medical Biometry and Medical Informatics, University of Freiburg, Germany
²Oral and Maxillofacial Oncology, Division of Maxillofacial Diagnostic & Surgical Sciences, Faculty of Dental Science, Kyushu University, Fukuoka, Japan
³Department of Pathology, University Hospital of Freiburg, Germany
⁴Department of Oral and Maxillofacial Surgery, University Hospital of Freiburg, Germany

Summary

Objectives: In this paper three statistical methods [logistic regression, classification and regression tree (CART), and fuzzy inference] for the prediction of lymph node metastasis in carcinoma of the tongue are compared. Methods: A retrospective collection of data in 75 patients treated for tongue cancer was carried out at the Clinic and Polyclinic for Oral and Maxillofacial Surgery at the University Hospital of Freiburg in Germany between January 1990 and December 1999; biopsy material was used for laboratory evaluations. Statistical methods for the prediction of lymph node metastasis were compared using ROC curves and accuracy rates. Results: All three methods show similar results for the prediction of lymph node metastasis with slightly superior results for fuzzy inference and CART. A great overlap is apparent in the ROC curves. The best result observed for fuzzy inference and CART was a sensitivity of 79.2% [95% confidence interval: (57.8%; 92.9%)] and a specificity of 86.3% (73.7%; 94.3%); the best result for predictions based on the logistic regression was a sensitivity of 66.7% (44.7%; 84.4%) and a specificity of 80.4% (66.9%; 90.2%). Accuracy rates of fuzzy method and CART were higher (accuracy rate for fuzzy method and CART: 84% (73.7%; 91.4%), for logistic regression method: 73.3%, 95%-CI: (61.9%; 82.9%)).

Conclusions: From a clinical point of view, the predictive ability of the three methods is not sufficiently large to justify use of these methods in daily practice. Other factors probably on the molecular level are needed for the prediction of lymph node metastasis.

Keywords

Tongue cancer, prediction of metastasis, logistic regression, classification and regression tree, fuzzy inference


Introduction

Cervical lymph node involvement is one of the most important determinants of prognosis in patients with squamous cell carcinoma of the oral cavity [1-3]. Because of the radicality of the definitive diagnostic and therapeutic approach it is desirable to have accurate predictions of cervical lymph node metastasis solely based on characteristics of the primary tumor. Many of those factors have already been investigated in numerous studies with controversial results. So far, no generally accepted prediction rule with sufficient accuracy could have been established. This may at least partially be explained by small study sizes, heterogeneity of patient populations with regard to treatment and prognosis and insufficient or inadequate statistical methodology.

Objectives

In this paper, we compare three different statistical approaches for the construction of a prediction rule in a relatively homogeneous population of patients with tongue cancer. In order to obtain reliable results from a study with moderate sample size we restrict ourselves to the investigation of only three potential factors. We will compare the resulting predictions by means of Receiver-Operating-Characteristic (ROC)-curves and accuracy rates.

Material

During the 10-year period from 1990 to 1999, 118 patients were treated for tongue cancer at the Clinic and Polyclinic for Oral and Maxillofacial Surgery, Albert Ludwig University, Freiburg, Germany. Patients whose initial biopsy material (hematoxylin and eosin staining) were of poor quality, whose follow-up period was less than 2 years, or who developed local recurrence at the primary site were excluded from the analyses. A total of 43 patients were excluded from the analysis for the following reasons: 6 patients didn’t have squamous cell carcinoma, in 1 patient a primary recurrence occurred within two years after surgery, 10 patients had a history of treatment of oral cancer, 16 patients were excluded due to missing information (including missing biopsy material), 7 patients were treated by radiation only, 3 patients had multiple primary oral cancers. Thus, 75 patients were included in the analysis. Descriptive information on these patients is shown in Table 1 and Figure 1.

Tumor size (as defined by the TNM classification, UICC 1998) was selected as a clinical factor, keratinization and mode of invasion were selected as histopathological factors involved in the occurrence of metastasis. Mode of invasion indicates the manner by which tumor cells infiltrate at the edge of the tumor proliferation area. This item has an important role in metastasis if mode of invasion increases, the pos-
sibility of metastasis increases. For keratinization, the percentage of cells showing any sign of keratinization was determined. Two co-authors (TN, DM) made histopathological assessments.

**Methods**

Three covariates (tumor size, mode of invasion and keratinization) were considered in the three models for the prediction of cervical lymph node metastasis. All three variables were treated as continuous variables.

The prediction rule for the fuzzy system was constructed on an independent dataset as described in [4]. Briefly, the method is based on a set of IF-THEN rules and a set of input and output membership functions with tumor size, mode of invasion and keratinization as input variables and the likeliness of lymph node metastasis as the outcome variable. A simple example for an IF-THEN rule is “IF mode of invasion is 1 THEN metastasis is unlikely”; this rule is actually included in Table 2: for a mode of invasion of 1 the response is negative irrespective of tumor size. Fuzziness is introduced by applying input and output membership functions to the IF-THEN rule. Note, output values of the fuzzy system do not easily transform into probability values.

Predictions for the logistic regression model [5, 6] were based on a regression model considering tumor size, mode of invasion and keratinization as potential covariates. Fractional polynomials, an extension of ordinary polynomials offering greater flexibility [7], were used to relax the assumption of linear covariate effects. After determining the functional form of the covariates, a stepwise-selection procedure using Akaike’s information criterion was applied starting with the full model including all three two-way interactions; bootstrap resampling with 1000 replications was used to assess the stability of the model selection process and ten-fold cross-validation was used to estimate accuracy rate, sensitivity and specificity [8].

Predictions of the three different methods were used to estimate receiver operating characteristic curves (ROC curves). Accuracy rate, sensitivity and specificity were calculated to assess the performance of the prediction rules; exact 95% confidence intervals were calculated according to the method described in [11]. Furthermore, a scatter plot matrix of predictions was evaluated.

**Results**

The fuzzy method resulted in the IF-THEN rule shown in Table 2. The predicted values range from -1 to 1 where larger values indicate the occurrence of lymph node metastasis.

### Table 1

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Mean (SD) or N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor size</td>
<td>T1: 32 (42.7%) T2: 29 (38.7%)</td>
</tr>
<tr>
<td>(TNM classification, UICC)</td>
<td>T3: 8 (10.7%) T4: 6 (8.0%)</td>
</tr>
<tr>
<td>Mode of invasion</td>
<td>3.37 (1.08)</td>
</tr>
<tr>
<td>Keratinization (in %)</td>
<td>36.6 (22.9)</td>
</tr>
</tbody>
</table>

Fig. 1 Scatterplot of tumor size and mode of invasion (Freiburg data)
None of the three covariates showed a non-linear effect in a multivariable fractional polynomial model, thus, all covariates were modelled as linear effects in the stepwise-selection. All two-way interactions and the covariate keratinization were excluded in the stepwise procedure resulting in the final logistic regression model shown in Table 3. These two covariates were included in almost all models selected in the 1000 bootstrap samples. Mode of invasion and tumor size both show a strong effect on the prediction of lymph node metastasis.

The CART method results in the classification tree depicted in Figure 2. With the CART method, tumor size is regarded as the most important factor on the prediction of lymph node metastasis. A tumor size greater than 2 is considered as an indication of lymph node metastasis. This group of patients is not subdivided again resulting in correct classification in 11 of 14 patients. The group of patients with a tumor size less than 2 is subdivided according to mode of invasion; patients with a mode of invasion larger than 4.4 have a higher risk of lymph node metastasis. The two resulting subgroups are not subdivided again. Overall, three subgroups are built: patients with a tumor size larger than 2, and patients with a tumor size of less than 2 and mode of invasion less/greater than 4.4. In the majority of bootstrap samples either tumor size greater than 2 or mode of invasion greater than 4.4 was selected as the first split criterion; keratinization was not considered as first split criterion in a single bootstrap sample.

ROC curves of the three methods are displayed in Figure 3. The CART method resulted in three different predictions (subgroups), thus, only two possible cutpoints exist resulting in a ROC curve with two steps. The fuzzy method and the logistic regression model resulted in a larger number of different predictions and thus the ROC curves for these methods have a larger number of steps. A visual inspection of the ROC curves reveals that the curves overlap to a large extent; none of the methods outperforms the others.

A single cutpoint is often chosen as the point closest to the upper left corner of the ROC curve. This approach results in a cutpoint of -0.35 for the logistic regression model and a cutpoint of -0.1 for the fuzzy method; for the CART method, patients with a tumor size less/or equal to 2 are predicted as free of lymph node metastasis. The corresponding contingency tables are shown in Tables 4 and 5; estimates of sensitivity, specificity and accuracy are based on ten-fold cross-validation for the logistic regression and CART method. Even though the same sensitivity and specificity values are observed for the fuzzy and the CART method, the prediction of the fuzzy and the CART method differ for two patients. The accuracy of the fuzzy and CART seems to be higher than the accuracy of the logistic regression method (logistic regression vs. Fuzzy: p = 0.011, logistic regression vs. CART: p = 0.021, evaluated by McNemar tests).

The scatterplot matrix displayed in Figure 4 shows the similarities and dissimilarities between the methods. Overall, predictions of the different methods are rather similar, e.g. the results of the fuzzy and CART method differ markedly only in two cases (middle row, left column).

Discussion

We have presented the results of three different statistical approaches for the construction of a prediction rule for cervical lymph node metastasis in patients with tongue cancer. In a series of 75 consecutive patients diagnosed and treated at the Clinic and Polyclinic for Oral and Maxillofacial Surgery of the University Hospital Freiburg, data on histopathological features of the primary tumor have been collected retrospectively. We have only considered
tumor size, mode of invasion and degree of keratinization as potential ingredients for a prediction rule. Tumor size and mode of invasion have been investigated in many other studies [3, 4, 12, 13] whereas degree of keratinization seems to have been rarely considered. To our knowledge, only in two small studies [14, 15] with 45 and 20 patients, respectively, a significant correlation between degree of keratinization and cervical lymph node metastasis has been found. In our study, this factor did not contribute to the predictive performance at all; thus it can be omitted from the prediction rules without compromising their accuracy.

With regard to the statistical approaches, logistic regression is the most widely used technique also in this setting [2, 3, 16]. It is closely related to linear discriminant analysis (used by [17]) or a variant according to [18]. The latter has been used by [4] and [12]. We used a two-step strategy in the model building process. In a first step, the technique of fractional polynomials was used to determine the functional form of the covariates. In a second step, a stepwise procedure, starting with a model with all main effects and two-way interactions, was used to select the final model. Bootstrap resampling was utilised in the second step to assess the stability of the model selection process. We avoided more general and flexible regression functions, as for example used in artificial neural networks, in order to get a perfect separation in the data set under consideration. These techniques generally tend to overfit the data and lack generalizability in independent data sets [19]. We are not aware that the classification and regression tree (CART) method has been applied in this setting. This approach, sometimes also called recursive partitioning analysis, leads to a very simple and easily understandable prediction rule. In order to avoid potential overfitting we used stringent splitting and stopping criteria [10, 20] based on corrected P-values for multiple testing. The prediction rule constructed with this method turned out to be based on tumor size and mode of invasion only.

We compared the logistic regression and the CART approach with the method based on fuzzy inference that has been sug-

**Table 4** Diagnostic result of logistic regression model based on ten-fold cross-validation (Accuracy: 73.3%, 95%-CI: [61.9%; 82.9%], Sensitivity: 58.3% [36.6%; 77.9%], Specificity: 80.4% [66.9%; 90.2%])

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Truth: Lymph node metastasis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Lymph node metastasis</td>
<td>14</td>
</tr>
<tr>
<td>No lymph node metastasis</td>
<td>10</td>
</tr>
</tbody>
</table>

**Table 5** Diagnostic result of fuzzy method and CART analysis, CART based on ten-fold cross-validation (Accuracy: 84.0%, 95%-CI: [73.7%; 91.4%], Sensitivity: 79.2% [57.8%; 92.9%], Specificity: 86.3% [73.7%; 94.3%])

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Truth: Lymph node metastasis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Lymph node metastasis</td>
<td>19</td>
</tr>
<tr>
<td>No lymph node metastasis</td>
<td>5</td>
</tr>
</tbody>
</table>

**Fig. 3** ROC curves of different methods
gested by [4]. In general, the three methods gave comparable results in terms of diagnostic accuracy. The ROC-curves, however, allow a much more detailed inspection of the predictive performance than a single value of sensitivity or specificity could provide. The best result was observed for CART and for fuzzy inference being slightly better than the best result obtained for logistic regression. The resulting observed sensitivities and specificities can be translated into a positive predictive value of 58.3% for logistic regression and 73.1% for CART and the fuzzy method and negative predictive values of 80.4% for logistic regression and 89.8% for CART and the fuzzy method, respectively. To base a definitive decision on such a prediction, its predictive performance still needs to be improved.

From Figure 1 it can be seen that this can not be achieved with more sophisticated statistical methods since there is a substantial overlap between patients with and without cervical lymph node metastasis. The only way towards a substantial improvement of the predictive performance leads through the identification of one or more powerful features probably on the molecular level. For that, adequate and efficient statistical methodology and careful study design is urgently needed. When scrolling through the literature we identified a number of publications where only univariate statistical analyses were performed (e.g. [13, 15, 21, 22]), or where basic requirements for prognostic studies [23] were not met. Especially small study sizes in contrast to a relatively large number of potential factors being investigated, as, for example, reported by [15], constitutes a serious problem. Thus there still seems to be sufficient room for the needed substantial improvements.

### Conclusions

Three statistical methods (logistic regression, classification and regression tree, and fuzzy inference) for the prediction of cervical lymph node metastasis in carcinoma of the tongue show similar results in a data set of 75 patients. The predictive ability of the three methods is not sufficiently large to justify use of these methods in daily practice; other factors are needed for the prediction of lymph node metastasis.

### References


Correspondence to:
Guido Schwarzer
Stefan-Meier-Strasse 26
D-79104 Freiburg
Germany
E-mail: sc@imbi.uni-freiburg.de