Application of Binary Response Regression Models to Detect Factors Influencing the Occurrence of Infection in Dental Surgeries

Polliyane Silva¹,a, Valdemiro Vigas²,b, Cristiane Silva¹,c, Ananda Barreto³,d, Paulo Oliveira Jr³,e, Taciana Savian¹,f, Isis Santos³,g

¹Departamento de Ciências Exatas, ESALQ, Universidade de São Paulo, Piracicaba, Brasil
²Instituto de Matemática, Universidade Federal de Mato Grosso do Sul, Campo Grande, Brasil
³Odontologia, Irmandade da Santa Casa de Misericórdia de Piracicaba, Piracicaba, Brasil

Abstract

Postoperative infection is common in dental surgery, for example, in the removal of the third molar. To control these and other postoperative complications, various studies have reported the use of some drug protocols, namely the prophylactic and the preemptive ones, using drugs such as dexamethasone and betamethasone. In this work, we used the generalized linear model via logistic regression to verify whether, in addition to the medications mentioned, some covariates that are frequently used in dental surgeries influence the occurrence, or not, of postoperative infection in surgeries for removal of the third molar. One of the main reasons that led us to employ such a model is because the response variable (having or not having infection) presents values of the binary type, in addition to being one of the most applied models in the area of health, among them, the dentistry area. The application of descriptive methods and analysis of

¹Ph.D. E-mail: vieira_polly@usp.br
²Ph.D. E-mail: valdemiro.vigas@ufms.br
³M.Sc. E-mail: cmrsilva@usp.br
⁴Specialist. E-mail: ananda_schlitter@hotmail.com
⁵M.Sc. E-mail: pauloafj@terra.com.br
⁶Ph.D. E-mail: tvsavian@usp.br
⁷M.Sc. E-mail: isisbohonaro@gmail.com
association via statistical tests were also used to choose other factors that influence the response variable infection in addition to medications. The AIC (Akaike Information Criterion) selection criterion, analysis of the difference in the deviations, and the analysis of residual using the half-normal plot for selection and the assumption of the proposed model were employed. The data set under analysis consists of 113 patients submitted to dental surgery in a specialized clinic in the city of Piracicaba, SP-Brazil between 2003 and 2018. Through the proposed model, some important information the covariates in relation to the patients submitted to dental surgery. One of the key information is that characteristics as Age and Dental extractions are associated with the inflammatory processes after surgery. This relationship indicates that the older the patient, the chance of having an infection after surgery increases. The analysis is similar to the Number of extractions.

Key words: Chi-square test; Dentistry surgeries; Generalized linear models; Half-normal plot; Logistic regression.

Resumen

La infección postoperatoria es común en cirugía dental, por ejemplo, en la extracción del tercer molar. Para el control de estas y otras complicaciones postoperatorias, diversos estudios han reportado el uso de algunos protocolos farmacológicos, a saber, el profiláctico y el preventivo, utilizando fármacos como la dexametasona y la betametasona. En este trabajo utilizamos el modelo lineal generalizado mediante regresión logística para verificar si, además de los medicamentos mencionados, algunas covariables que se utilizan con frecuencia en las cirugías dentales influyen en la aparición o no de infección posoperatoria en las cirugías de extracción del tercer tercio molar. Una de las principales razones que nos llevó a emplear dicho modelo es porque la variable respuesta (tener o no tener infección) presenta valores de tipo binario, además de ser uno de los modelos más aplicados en el área de la salud, entre ellos, el área de odontología. También se utilizó la aplicación de métodos descriptivos y análisis de asociación a través de pruebas estadísticas para elegir otros factores que influyen en la variable respuesta infección además de los medicamentos. Se empleó el criterio de selección AIC (Akaike Information Criterion), el análisis de la diferencia en las desviaciones y el análisis de residuos utilizando la seminormal para la selección y la asunción del modelo propuesto. El conjunto de datos bajo análisis consta de 113 pacientes sometidos a cirugía dental en una clínica especializada en la ciudad de Piracicaba, SP-Brasil entre 2003 y 2018. A través del modelo propuesto, algunas informaciones importantes las covariables en relación a los pacientes sometidos a cirugía dental. Uno de los datos clave es que características como la edad y las extracciones dentales están asociadas a los procesos inflamatorios posteriores a la cirugía. Esta relación indica que a mayor edad del paciente, aumenta la probabilidad de tener una infección después de la cirugía. El análisis es similar al Número de extracciones.

Palabras clave: Cirugías de odontología; Gráfico half normal; Modelos lineales generalizados; Prueba de chi-cuadrado; Regresión logística.
1. Introduction

Lower third molar extraction is the most commonly procedure performed by
surgeons maxillofacial all around the world (Cankaya et al., 2011). Like any type
of surgery, it has risks, accidents, and complications that can affect the patients
(Azenha et al., 2014; Chiapasco et al., 1993). According to Alsaedi (2021), tooth
extraction was a significant factor when associated with tooth decay, which is a
type of tooth infection. Santosh (2015) claims that one of the relevant factors for
the occurrence of complications associated with tooth removal is the patient’s age.

Among the most frequent postoperative complications in dentistry studies most
common related to third molar removal, the alveolar osteitis and the surgical field
infection stand out (Bui et al., 2003) which in most situations make this type
of procedure an uncomfortable experience for patients who are submitted to it.
Therefore, efficient control of inflammation performs an important role once the
prophylactic and preemptive effects of various drugs have been investigated and
the results have largely supported the use of corticosteroids (Mehra et al., 2013;
Sotto-Maior et al., 2011).

The literature presents a linear association between the duration of the
procedure and the incidence of infection Guglielmo et al. (1983); Capuzzi et al.
(1994); Lacasa et al. (2007). From this perspective, Lacasa et al. (2007) affirms
that the difficulty is characterized by the need for osteotomy during the surgical
procedure and can increase the incidence of infections from 3.5% to 12.7%.
However, the incidence of localized infections of the oral cavity decreased with
the use of antibiotic prophylaxis when compared to placebo in patients submitted
to tooth extraction (Moreno-Drada & Garcia-Perdomo, 2016; Susarla et al., 2011).

The indiscriminate use of antibiotics increases the risk of toxicity related to
allergic reactions, secondary infections and bacterial resistance (Poeschl et al.,
2004; Zeitler, 1995). The prevalence of adverse events after antibiotic therapy
varies from 6 to 7% (Ataoglu et al., 2008), where the main events are nausea and diarrhea
(Limeres et al., 2009). Recommendations for the use of antibiotic prophylaxis in surgery is to apply the antimicrobial in the correct moment and as
soon as possible. Furthermore, they must be efficient and present low toxicity.
Thus, the use of prophylaxis should be considered and indicated for patients
submitted to more invasive surgeries in which osteotomies are necessary for
one or more operative sites, or in patients with pre-existing diseases, or even
immunocompromised (Laskin, 2003; Nishimura et al., 2008).

In an attempt to control postoperative complications after third molar
extraction, numerous studies have reported the effectiveness of some drugs such
as dexamethasone and betamethasone (Chaudhary et al., 2015; Lima et al., 2018;
Vicentini et al., 2018). Both have lower mineralocorticoid activity, a half-life of
36-72 hours, and 25 times greater effectiveness than hydrocortisone in reducing
inflammatory mediators that have less effect on leukocyte chemotaxis. Also, in
programmed elective surgeries, there is a unique opportunity to prepare patients
for the surgical procedure, effectively preventing the triad of pain, edema, and
infection, symptoms of classic occurrences of invasive surgical procedures. This
preventive care provide these patients a more comfortable postoperative period, free from complications secondary to surgery, and for this, the use of medications is inevitable.

In many situations in the dental field, statistical tools are used to ease the interpretation of data and assist in decision-making. Among the statistical methodologies used in this area, logistic regression stands out because in many cases the researcher is interested in analyzing data whose response variable is of the binary type. Many works contributed to the modeling of dental data. For example, Alsaedi (2021) used logistic regression to determine some dependent variables that were effective on dental caries.

Thus, the main objective of this manuscript is to verify the effectiveness of drug protocols (prophylactic and preemptive), to verify some characteristics collected from patients undergoing surgery, in addition to factors related to surgical procedures. Through the analysis of generalized linear models, a verification on which of these factors influence postoperative infection is presented. The quantitative analysis by descriptive and inferential methods over through the chi-square test further of the generalized linear models via logistic regression were used with the support of the R software, bringing complementary information and the discussion of results.

2. Materials and Methods

The data used are referring to 113 patients that were submitted to dental surgery in a specialized clinic in the city of Piracicaba between 2003 and 2018. The preoperative protocols used are prophylactic and preemptive, which are described below:

i. **Prophylactic protocol**: 1 gram of Cephalexin was used 2 hours before surgery and 1 gram 6 hours after the first dose. In the corticoid regime, two possibilities were considered: intramuscular injectable (IM) use of 1 ampoule of Celestone soluspan, 15 minutes before surgery, or the use of 8 or 12 mg doses of dexamethasone via oral (VO), 2 hours before surgery. Patients were also instructed to realize yawn with 0.12% chlorhexidine starting 1 or 2 days before surgery. For the postoperative period, anti-inflammatory and analgesics were recommended, if necessary.

ii. **Preemptive protocol**: 500 mg cefadroxil was used starting of prophylactic form with 1 gram, 2 hours before surgery and maintained for every 12h for up to 5 days. In the corticoid regime, two possibilities were considered: intramuscular injectable (IM) use of 1 ampoule of Celestone soluspan, 15 minutes before surgery, or the use of 8 or 12 mg doses of dexamethasone via oral (VO), 2 hours before surgery.

The surgeries were performed by the same surgeon, following rigid asepsis and antisepsis standards and using sterile gloves. The drugs and antibiotics used were
chosen according to the rules for the use of prophylactic antibiotics, observing their effective spectrum of action against the germs commonly present in the oral cavity. Furthermore, using first-generation cephalosporins, cephalexin in prophylaxis, due to its short dosage period, and cefadroxil 500 mg in preemptive treatment, as recommended by the Ministry of Health regarding the use of prophylactic antibiotics. The predominant factor for indicating a protocol is associated with the degree of difficulty of the surgery or the number of elements to be extracted in a single session, considering the surgical time and magnitude of the surgical trauma. In patients with a history of allergy to cephalosporin or penicillin, who presented a beta-lactam ring, clindamycin 300 mg was used.

Surgical sequence values were: 38, 28, 48, and 18, as shown in Figure 1, when the 4 elements are extracted in a single session by standardization. In the surgeries, multiplier pieces with external irrigation without the use of high rotation were used, and all the principles of intra and extra oral asepsis and antisepsis were followed, with the use of sterile fields.

![Image representing the number of teeth.](image)

**Figure 1:** Image representing the number of teeth.

The characteristics of the patients were recorded. The variable of interest (response variable) \( Y \): **Postoperative infection** (0 = no, 1 = yes) is presented, in addition to the covariates:

- \( X_1 \): **Gender** (0 = female, 1 = male);
- \( X_2 \): **Age** (in years);
- \( X_3 \): **Number of extracted teeth** (1, 2, 3, 4, and > 4);
- \( X_4 \): **Type of preoperative protocol** (0 = prophylactic, 1 = preemptive);
- \( X_5 \): **Surgery Complication** (0 = without complications, 1 = with complications);
- \( X_6 \): **Surgery difficulty degree** (0 = low, 1 = medium, 2 = high);
- \( X_7 \): **Type of drug treatment** (0 = celestone, 1 = dexamethasone).
Based on these covariates we can evaluate the effectiveness of protocols; to verify the complications according to the difficulty/surgical time; and as the use of treatments applied to patients undergoing surgery. Also, it was observed how these characteristics influence the variable of interest, i.e., the occurrence of infections.

Descriptive analyses were carried out to obtain more information about the data. Subsequently, the chi-square test of independence was used, because the interest was to study the relationship between the covariates, two by two, following their assumptions.

As there is an interest in describing some existing association between the covariates and the variable response infection, the regression model was chosen, in particular, the generalized linear model (GLM) via logistic regression, because the response variable is of binary type with more than one covariate. The command glm is available in R Software for the adjustment of the generalized linear model to binary data.

The response variable used Postoperative infection is dichotomous with the levels: with infection dental \( y_i = 1 \) and no infection \( y_i = 0 \) for each patient \( i = 1, \ldots, 113 \). In this context, we are considering that \( Y_i \sim \text{Bernoulli} (\pi_i) \); where \( \pi_i \) is the probability that the \( i_{th} \) patient presents the response under investigation, i.e., \( P(Y_i = 1) = \pi_i \) for \( i = 1, \ldots, 113 \). The systematic component represents the information to the researcher about the phenomenon. This information is translated by the combination of parameters and covariates expressed by:

\[
\eta = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p,
\]

where: \( \eta \) indicates the linear predictor; \( X_i \) is the covariates vector of the design matrix \( X \); \( \eta = (\beta_0, \beta_1, \ldots, \beta_k)^T \) is the vector of unknown parameters to be estimated and \( p \) represents the number of parameters.

As the Bernoulli distribution was assumed as the response variable, the link functions \( g^{-1}(\pi) \) used were: probit, logit, and complementary log-log. The use of the three link functions is justified by the unbalance in the proportion of patients about the infection response variable. However, in this case for didactic purposes, only the logistic model was presented.

The logistic regression is one of the best-known methods and is applied in several knowledge areas, among them Biology, as analyzed in Javali & Pandit (2012), which verified the dependence of oral health diseases, dental caries, and periodontal disease through the application of the logistic regression model.

Starting from the fact that \( Y_i \) is a binary response variable, we assume the following logistic regression model to associate the covariates \( X_i \) with the response variable, so that:

\[
Y_i \sim \text{Bernoulli} (\pi_i); \quad i = 1, 2, \ldots, 113.
\]

\[
\pi_i = E(Y_i | X_i) = \frac{\exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p)}.
\]

Using the logit link function we have:

\[
\text{logit}(\pi_i) = \log \left( \frac{\pi_i}{1 - \pi_i} \right) = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_p X_{pi}.
\]
The command `family = binomial(link = 'logit')` is available in R Software for the adjustment of the generalized linear model to binary data with the logit link function.

To indicate which model is more suitable for the data, the Akaike information criterion (AIC) was introduced. Also, the deviation analysis was used to test the significance of the inclusion of new covariates in the embedded models or their possible interactions, in order to verify which of these is the most appropriate for the data set under study. The `anova` command of R Software provides a table of deviation analysis. More information in Nelder & Wedderburn (1972), Cordeiro & Neto (2004), Cordeiro & Demétrio (1986) and Moral et al. (2017).

An important step after modeling is the residual analysis because from this it is possible to verify the assumptions of the proposed model. According to the author Cordeiro & Neto (2004), the residuals in the GLM’s are used to explore the suitability of the fitted model concerning the link function and linear predictor terms in addition to identifying outliers. For the residuals analysis in this work, the normal graph of probabilities with simulation envelope was used (Moral et al., 2017) via the command `hnp`.

3. Results and Discussion

To obtain more information about the data set under study, a descriptive analysis was initially performed, being the results summarized in Tables 1 and 2 and in Figures 2 and 3. It was found that from the 113 patients under study, 95 (84%) had no infection, while 18 patients (16%) had the infection, evidencing an imbalance in the proportion of these patients about the variable of interest. Furthermore, it was observed that the relative frequencies of the covariates studied in Table 1 have mostly close values, initially indicating descriptively that these covariates do not have distinct results for patients with or without infection.

Regarding the covariate patient’s age, it was observed that the measures of central tendency such as the mean, median, quartiles, and of variability such as the coefficient of variation (CV), showed different behaviors for the patients with and without infection, being also observed that the age distribution represented by Figure 2 shonwn a distinct behavior with atypical observations for both types of patients.

As for the covariate number of teeth extracted, most patients removed four teeth regardless of the presence of infection or not. Although the number of teeth extracted was higher in individuals who did not have an infection, which corresponds to a frequency of 54 observations when compared to patients with infection who have a frequency of 11 observations, in percentage terms the results become more similar, that is, the percentage of individuals without infection and with infection who extracted four teeth correspond to 56.84% and 61.12%, respectively, as presented in Figure 3.
Table 1: Frequency distribution of some variables under study.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Infection</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No (n = 95)</td>
<td>Yes (n = 18)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>44 (46.31%)</td>
<td>10 (55.55%)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>51 (53.68%)</td>
<td>8 (44.44%)</td>
<td></td>
</tr>
<tr>
<td>Type of Protocol Preoperative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prophylactic</td>
<td>62 (65.26%)</td>
<td>10 (55.55%)</td>
<td></td>
</tr>
<tr>
<td>Preemptive</td>
<td>33 (34.74%)</td>
<td>8 (44.44%)</td>
<td></td>
</tr>
<tr>
<td>Complication in surgery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>67 (70.53%)</td>
<td>15 (83.33%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>28 (29.47%)</td>
<td>3 (16.66%)</td>
<td></td>
</tr>
<tr>
<td>Degree of difficulty of the surgery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>15 (15.79%)</td>
<td>2 (11.11%)</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>22 (23.16%)</td>
<td>4 (22.22%)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>58 (61.05%)</td>
<td>12 (66.66%)</td>
<td></td>
</tr>
<tr>
<td>Type of drug treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Celestone</td>
<td>49 (51.58%)</td>
<td>10 (55.55%)</td>
<td></td>
</tr>
<tr>
<td>Dexamethasone</td>
<td>46 (48.42%)</td>
<td>8 (44.44%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Descriptive measures regarding the age of patients with and without infection

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>1st Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Quartile</th>
<th>Max</th>
<th>CV(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>With infection</td>
<td>14.00</td>
<td>21.25</td>
<td>24.00</td>
<td>26.89</td>
<td>30.75</td>
<td>48.00</td>
<td>36.63</td>
</tr>
<tr>
<td>No infection</td>
<td>15.00</td>
<td>18.00</td>
<td>21.00</td>
<td>23.78</td>
<td>26.00</td>
<td>56.00</td>
<td>33.72</td>
</tr>
</tbody>
</table>

Figure 2: Boxplot of age for the infection variable.
An Application of Binary Response Regression Models

Table 3 presents an analysis to verify the existence of any association between the covariates using the chi-square test. It was observed that the Number of teeth extracted \( (X_3) \) has an association with the Degree of difficulty of surgery \( (X_6) \); Type of preoperative protocol \( (X_4) \) has an association with the Degree of difficulty of surgery \( (X_6) \) and Type of drug treatment \( (X_7) \), and that Degree of difficulty of surgery \( (X_6) \) has a relationship with Type of drug treatment \( (X_7) \). The reason for this association is justified by the fact that the p-value was less than the adopted \( \alpha \) significance level of 10%.

Table 3: Matrix of the covariates studied with their respective associated p-values.

<table>
<thead>
<tr>
<th>Variables</th>
<th>( X_1 )</th>
<th>( X_3 )</th>
<th>( X_4 )</th>
<th>( X_5 )</th>
<th>( X_6 )</th>
<th>( X_7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_1 )</td>
<td>-</td>
<td>0.41</td>
<td>0.67</td>
<td>0.48</td>
<td>0.83</td>
<td>0.21</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>0.41</td>
<td>-</td>
<td>0.33</td>
<td>0.65</td>
<td>0.03</td>
<td>0.21</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>0.67</td>
<td>0.33</td>
<td>-</td>
<td>0.91</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>0.48</td>
<td>0.65</td>
<td>0.91</td>
<td>-</td>
<td>0.24</td>
<td>0.07</td>
</tr>
<tr>
<td>( X_6 )</td>
<td>0.83</td>
<td>0.03</td>
<td>0.01</td>
<td>0.24</td>
<td>-</td>
<td>0.02</td>
</tr>
<tr>
<td>( X_7 )</td>
<td>0.21</td>
<td>0.21</td>
<td>0.01</td>
<td>0.07</td>
<td>0.02</td>
<td>-</td>
</tr>
</tbody>
</table>

According to the researchers’ guidelines, it was defined that the variable degree of difficulty \( (X_6) \) would be removed from the analysis. Therefore, the covariates used in the analysis were: Gender \( (X_1) \); Age \( (X_2) \); Number of teeth extracted \( (X_3) \); Type of preoperative protocol \( (X_4) \); Complication during surgery \( (X_5) \) and Type of drug treatment \( (X_7) \).

Since the results of the probit, logit, and complementary log-log link functions were similar, it was decided to use the logistic link function under the following structure:

\[
\pi_i = \log \left( \frac{\pi_i}{1 - \pi_i} \right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i};
\]

where \( i = 1, 2, \ldots, 113 \).
The estimated values with their respective $p$-values of the model with logit link function were presented in Table 4 and through this, it was observed that besides the intercept, the covariates: Age; the Number of teeth extracted, and Complication in surgery were relevant at the 10\% level of significance.

Table 4: Estimates of the parameters of the logistic regression models.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimates</th>
<th>$SE$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>-5.26</td>
<td>1.74</td>
<td>0.01</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.34</td>
<td>0.56</td>
<td>0.51</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.08</td>
<td>0.07</td>
<td>0.03</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.53</td>
<td>0.30</td>
<td>0.07</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-0.61</td>
<td>0.72</td>
<td>0.40</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>-1.31</td>
<td>0.76</td>
<td>0.08</td>
</tr>
<tr>
<td>$\beta_6$</td>
<td>0.54</td>
<td>0.74</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Based on the fact that the covariates Age, Number of teeth extracted and surgical complication were relevant, we considered some submodels individually and their possible combinations to verify if they fit the data well result 10 candidate models $M_j; j = 1, 2, \ldots, 10$. As the 10 candidate models are fitted, the difference analysis of the deviations was used to verify whether the proposed models fits the data. Tables 5, 6 and 7 present the results regarding three fitted model configurations using the logit link function. Through the tables, we can see that models 4 and 10 initially show satisfactory evidence of dental data, as the candidate models presented a $p$-value lower than 10\%.

Table 5: Values of degree of freedom numbers with their differences, standard deviations with their differences and $p$-value, using the logit link function in configuration 1.

<table>
<thead>
<tr>
<th>Candidate Models</th>
<th>g.l.</th>
<th>Deviations</th>
<th>Difference df</th>
<th>Diff. deviations</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_2$: Number of teeth</td>
<td>111</td>
<td>97.74</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$M_1$: Age</td>
<td>111</td>
<td>97.18</td>
<td>0</td>
<td>0.49</td>
<td>-</td>
</tr>
<tr>
<td>$M_4$: Age + Number of teeth</td>
<td>110</td>
<td>93.12</td>
<td>1</td>
<td>4.06</td>
<td>0.04</td>
</tr>
<tr>
<td>$M_5$: Age*Number of teeth</td>
<td>109</td>
<td>93.12</td>
<td>1</td>
<td>0.01</td>
<td>0.96</td>
</tr>
<tr>
<td>$M_{10}$: Number of teeth<em>Complication</em>Age</td>
<td>105</td>
<td>84.00</td>
<td>4</td>
<td>8.12</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 6: Values of degree of freedom numbers with their differences, standard deviations with their differences and $p$-value, using the logit link function in configuration 2.

<table>
<thead>
<tr>
<th>Candidate Models</th>
<th>g.l.</th>
<th>Deviations</th>
<th>Difference df</th>
<th>Diff. deviations</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_3$: Complication</td>
<td>111</td>
<td>97.74</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$M_1$: Age</td>
<td>111</td>
<td>97.18</td>
<td>0</td>
<td>0.56</td>
<td>-</td>
</tr>
<tr>
<td>$M_6$: Age + Complication</td>
<td>110</td>
<td>94.51</td>
<td>1</td>
<td>2.67</td>
<td>0.10</td>
</tr>
<tr>
<td>$M_7$: Age*Complication</td>
<td>109</td>
<td>94.50</td>
<td>1</td>
<td>0.02</td>
<td>0.89</td>
</tr>
<tr>
<td>$M_{10}$: Number of teeth<em>Complication</em>Age</td>
<td>105</td>
<td>85.01</td>
<td>4</td>
<td>9.40</td>
<td>0.05</td>
</tr>
</tbody>
</table>
An Application of Binary Response Regression Models

Table 7: Values of degree of freedom numbers with their differences, standard deviations with their differences and p-value, using the logit link function in configuration 3.

<table>
<thead>
<tr>
<th>Candidate Models</th>
<th>g.l.</th>
<th>Deviations</th>
<th>Difference df</th>
<th>Diff deviations</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>M&lt;sub&gt;3&lt;/sub&gt;: Complication</td>
<td>111</td>
<td>97.74</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M&lt;sub&gt;2&lt;/sub&gt;: Number of teeth</td>
<td>111</td>
<td>97.67</td>
<td>0</td>
<td>0.07</td>
<td>-</td>
</tr>
<tr>
<td>M&lt;sub&gt;8&lt;/sub&gt;: Number of teeth+Complication</td>
<td>110</td>
<td>96.58</td>
<td>0</td>
<td>1.00</td>
<td>0.29</td>
</tr>
<tr>
<td>M&lt;sub&gt;9&lt;/sub&gt;: Number of teeth*Complication</td>
<td>109</td>
<td>96.02</td>
<td>1</td>
<td>0.55</td>
<td>0.45</td>
</tr>
<tr>
<td>M&lt;sub&gt;10&lt;/sub&gt;: Number of teeth<em>Complication</em>Age</td>
<td>105</td>
<td>85.00</td>
<td>4</td>
<td>12.03</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Then the parameter estimates were estimated again with only the significant covariates represented by the Table 8 and it was observed that the covariates age and number of extracted teeth were relevant at the significance level of 10%. Figure 4 present the normal probability plot with simulation envelope for the model indicating a good fit.

Table 8: Estimate of the parameters of the regression model with the covariates Age and Number of extracted teeth without the interaction.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimates</th>
<th>EP</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>β&lt;sub&gt;0&lt;/sub&gt;</td>
<td>-5.36</td>
<td>1.65</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>β&lt;sub&gt;1&lt;/sub&gt;</td>
<td>0.07</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>β&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0.56</td>
<td>0.29</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Figure 4: Normal probability plot with simulated envelope for the final model.

Therefore the final regression model for binary data is represented in the form:

\[
\text{logit}(\pi_i) = \log \left( \frac{\pi_i}{1-\pi_i} \right) = -5.36 + 0.07 \text{ Age}_i + 0.56 \text{ Number of extracted teeth}_i,
\]

\(i = 1, 2, \ldots, 113\).

Considering the use of the logistic model, some interpretations can be made on the estimated parameters, such as the odds ratio. Therefore, for the variables Age and Number of extracted teeth, the odds ratios are interpreted as follows: the odds of getting an infection for each year of age is \(\exp(0.07) = 1.07\); if there was interest in a larger interval in the difference, for example, 4 years, in that case, for each increase in 4 years, the chance of getting the infection is \(\exp(4 \times 0.07) = 1.32\) and the odds of getting an infection for each extracted tooth is \(\exp(0.56) = 1.76\).
4. Conclusions

In this study, the generalized linear model via logistic regression was used to verify whether, in addition to the medication protocols presented, some covariates frequently used in dental surgeries influence the response variable, occurrence or not of infection. Through the methodologies used, it was observed that only the variables age and number of extracted teeth were significant about the variable infection. This result corroborates other research in the area (Santosh, 2015; Alsaedi, 2021), in which the patient’s age and the number of extracted teeth have influence on inflammatory processes after surgery and occurrence of infection. The logit link function was chosen due to the greater ease of interpreting the parameters and because of its recurrent use in scientific articles in the dental field.

Acknowledgements

We thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the financial support for this research.

References


