Quality of logistic regression reporting in studies of tooth survival after periodontal treatment


Abstract

Aim: To evaluate the quality of reporting of logistic regression models used to assess risk factors for tooth loss in patients who have received periodontal treatment.

Materials/methods: The PubMed, EMBASE, BIOSIS Citation Index, CINAHL, Web of Science, and LILACS electronic databases were searched up to 01 March 2014 to identify interventional longitudinal studies assessing risk factors for tooth loss after periodontal treatment. The reference lists of included studies were searched manually. No language restriction was applied to the search. Quality of reporting of logistic regression models was assessed using analytical and documentation criteria with a 15-item checklist. Criteria were judged as met (adequately reported) or not met (not reported). All searches, selection, data extraction, and quality assessment were performed independently and in duplicate.

Results: Of 621 records initially retrieved, 24 articles were included in the analysis. Less than 30% of all 360 datapoints were met. “Coding of independent variables” was reported most frequently [n = 22 (83%) articles]. Criteria such as “internal and external validation of the model” were not met in any study assessed.

Conclusion: The reporting of logistic regression models in studies assessing risk factors for tooth loss in patients who have received periodontal treatment is not optimal.

Key words: logistic regression; methodological quality; periodontal treatment; prognostic; risk factors; tooth survival

Accepted for publication 22 September 2014

Conflict of interest and source of funding statement

The authors declare that they have no conflict of interests. This study was self-funded by the authors.

Risk factors for tooth loss have been researched thoroughly in recent years (Chambrone et al. 2010). Many factors have been identified, and this knowledge is vital for determining the prognosis of a tooth affected by periodontal disease (Faggion et al. 2007). To provide accurate information, however, the reliability of risk factor assessment is a sine qua non condition. Such assessment is usually conducted using logistic regression methods when outcome variables are nominal or categorical (i.e. yes/no categorization of tooth loss) (Nathanson & Higgins 2008). These methods are often used to test associations among several variables (e.g. risk factors) and determine the probability of outcomes (e.g. tooth loss). Although logistic regression is a widely used statistical procedure, the medical literature has demonstrated the existence of associated analytic and reporting problems (Khan et al. 1999, Moss et al. 2003, Mikolajczyk et al. 2008, Kalil et al. 2010, Bagley et al. 2001).

Because there is a great variety of logistic regression approaches
Logistic regression differs from linear regression methodologies (Kleinbaum & Klein 2010), the use of inaccurate regression methods may generate unreliable results that directly interfere in the decision-making process, as for instance, in the case of improper use of conventional logistic regression analysis for the assessment of site-based instead of patient-based outcomes. For clustered data, multilevel analysis or similar approaches should be used instead. To date, no study has examined the quality of reporting of logistic regression analysis methods in studies of tooth loss. Explicit reporting is necessary to allow other researchers to replicate results (Bagley et al. 2011) and to improve interested readers’ understanding of the methodology.

The objective of this study was to evaluate the reporting of logistic regression methods in a systematic review of studies assessing risk factors for tooth loss after periodontal treatment. Moreover, we provide several recommendations to assist researchers in better reporting of these statistical assessments in studies of clinical periodontal research.

Materials and Methods

Logistic regression model definition

A logistic regression model is used frequently to assess the association between a binary variable and a disease (Prentice 1976). For example, the present work evaluated studies considering the binary variable of tooth loss/survival, predicted by several independent variables, such as tooth mobility, increased pocket depth (PD), and smoking.

In the logistic regression model, the logit transformation of the probability of tooth loss (π) is estimated as a linear function of p covariates (Khan et al. 1999):

\[
\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p, \tag{1}
\]

where \(\beta_0\) is the intercept, and \(\beta_1\) to \(\beta_p\) are the regression coefficients for covariates \(x_1\) to \(x_p\) respectively. Logistic regression differs from linear regression in two main ways (Stoltzfus 2011): (1) logistic regression aims to estimate the probability that the outcome will be observed, with logit transformation ensuring that the range of estimated probability is always 0 to 1; and (2) in logistic regression, the interpretation of a regression coefficient for a covariate, such as \(\beta\) in equation 1, is performed using the log odds or log odds ratio. Thus, the exponentiated coefficients are more interpretable and should be reported.

Literature search

A comprehensive literature search was conducted using the PubMed, EMBASE, BIOSIS Citation Index, CINAHL, Web of Science, and LILACS electronic databases, with the last search performed on 1 March 2014. The literature search strategy and key words used are reported in Table 1. The reference lists of included papers were also searched manually to identify additional potentially eligible studies. Finally, the grey literature was searched using OpenSIGLE (System for Information on Grey Literature in Europe). Manual searching included a complete consideration of articles published in the Journal of Periodontology, Journal of Clinical Periodontology, and Journal of Periodontology Research through March 2014. Two authors (CMF and LC) performed all literature searches independently and in duplicate. Disagreements about study selection were resolved by discussion until consensus was achieved. No language restriction was applied to the literature search.

Inclusion and exclusion criteria

Full-length journal articles of prospective and retrospective interventional studies that assessed tooth loss/survival after conservative (i.e. root scaling) or surgical (i.e. any form of periodontal surgery), and supportive periodontal therapy (SPT) using logistic regression analysis were included in this study. To be included, the study should report a logistic regression model evaluating the influence of variables on any form of tooth survival. Studies that used outcome measures other than tooth loss/survival, such as changes in PD and clinical attachment level; those reporting the use of statistical approaches other than logistic regression; and those in which no active periodontal treatment was performed were excluded from this review. We did not use the duration of follow-up as a criterion for study inclusion.

Data selection and extraction

Two reviewers (CMF and LC) independently screened the titles and abstracts of studies retrieved from literature searches. In this phase, articles that did not meet the inclusion criteria were excluded and the reasons for exclusion were recorded. Duplicate publications were also excluded. When titles and abstracts provided insufficient information to evaluate study inclusion, final decisions were based on examination of the full texts of articles. Disagreements about the inclusion of studies in the analysis were resolved by discussion until consensus was achieved.

Data, including those describing the main characteristics of included studies (e.g. methodology, independent variables), were extracted directly into a standardized table. Data entry was coded after the assessment of full texts of included articles into an Excel spreadsheet (1: criterion met, 2: unclear, and 0: criterion not met).

Two authors (CMF and LC) performed data extraction independently and in duplicate. Disagreements in this phase were resolved by discussion until consensus was achieved.

Assessment of reporting of logistic regression methodology

We used the methodology proposed by Bagley et al. (2011) and Kalil et al. (2010) to assess the quality...
of reporting of logistic regression analysis methods in included studies. This methodology utilizes analytical and documentation criteria (Table 2) and a YES (criterion adequately met)/NO (criterion not met)/UNCLEAR (insufficient information available for accurate assessment) structure. We also assessed whether authors included the differences in patients’ follow-up lengths in the logistic regression model. Two reviewers (CMF and LC) conducted a pilot test of the criteria using three papers reporting logistic regression methodology in periodontology. These three papers were not included in the final sample. An experienced statistician (YKT) supervised the preliminary assessment, and the form was refined as necessary. CMF and LC then assessed reporting in the included studies independently and in duplicate. YKT cross-checked the entries to ensure precision and resolved disagreements during the quality assessment.

Data were descriptively reported in absolute numbers and percentages of criteria met. Furthermore, quality of reporting was analysed by grouping publications by time periods.

### Results

#### Articles included in the analysis


#### Study characteristics

All included studies had longitudinal retrospective designs and were published between 1998 and 2014. Studies were conducted in 11 different periods.

### Table 2. Criteria for the methodological assessment of logistic regression analysis used in studies related to tooth loss in periodontally treated patients

<table>
<thead>
<tr>
<th>Analytical criteria</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Presence of linearity for continuous or transformed categorical variables: Inadequate: no description of the linearity assumption. Adequate: description of the linearity assumption. Preferably some kind of justification is provided, e.g. prior clinical knowledge, exploratory plot of the data.</td>
<td></td>
</tr>
<tr>
<td>3. Test for interaction for independent variables: Inadequate: use of interaction term without providing its statistical or clinical justification. Adequate: use of interaction term providing its statistical or clinical justification.</td>
<td></td>
</tr>
<tr>
<td>4. Main effect estimates, 95%CI, and statistical significance: Inadequate: odds ratio and 95%CI (or $p$ values) for each variable included in the final model is not provided. Adequate: odds ratio and 95%CI (or $p$ values) for each variable included in the final model is provided.</td>
<td></td>
</tr>
<tr>
<td>5. Prevention of overfitting: Inadequate: The ratio between the number of the less common outcomes and the number of independent variable parameters (“events”) is less than 10. Adequate: the ratio is equal or greater than 10.</td>
<td></td>
</tr>
<tr>
<td>6. Report of outlying data: Inadequate: If present, no report about how these data were treated. Adequate: If present, report about how these data were treated.</td>
<td></td>
</tr>
<tr>
<td>7. Goodness-of-fit procedures: Inadequate: no reporting of goodness-of-fit test (e.g. Hosmer-Lemeshow test) or regression diagnostics. Adequate: reporting of goodness-of-fit test (e.g. Hosmer-Lemeshow test) or regression diagnostics. Area under ROC curve should be provided if prediction is of interest.</td>
<td></td>
</tr>
<tr>
<td>8. Multicolinearity of predictive models: Inadequate: If high correlation between two independent variables is discussed in the study, but a test for colinearity is not provided. Adequate: If high correlation between two independent variables is discussed in the study and a test for colinearity is provided.</td>
<td></td>
</tr>
<tr>
<td>9. Internal validation of predictive models: Inadequate: no data splitting (including cross-validation), or bootstrapping, or jackknife. Adequate: provision of at least one of these three procedures.</td>
<td></td>
</tr>
<tr>
<td>11. Evidence for clinical decision rules: Level 1: At least one prospective validation in a different population and one impact analysis, demonstrating change in clinician behaviour and improve patient outcomes. Level 2: At least one prospective validation in a broad spectrum of patients and clinicians, or validated in several smaller settings that differ from one another. Level 3: At least one prospective validation in a narrow spectrum of patients, similar to those evaluated in the derivation sample. Level 4: Just internal validation (data splitting, bootstrapping, jackknife) from retrospective data. Level 5: Derivation only. No internal validation.</td>
<td></td>
</tr>
<tr>
<td>Documentation criteria</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>12. Selection of independent variables: Inadequate: no provision of the rationale for including each variable. Adequate: provision of the rationale for including each variable, e.g. prior studies/clinical knowledge, exploratory analysis.</td>
<td></td>
</tr>
<tr>
<td>13. Coding of independent variables: Inadequate: if variables included in the model cannot have coding identifiable by our team. Adequate: if variables included in the model have coding identifiable by our team.</td>
<td></td>
</tr>
<tr>
<td>14. Fitting procedure: Inadequate: no provision of how the final model was selected. Adequate: provision of model selection method used, such as automatic procedures (e.g. forward inclusion or backward elimination or stepwise or best subset, preferably with description of the criteria used), non-automated backward selection, or prior knowledge.</td>
<td></td>
</tr>
</tbody>
</table>
countries. The mean number of patients included in logistic regression analyses was 203 (range, 25–776; total, 4878 patients; Table 3). Most patients included in the studies had an average age of more than 40 years old. The studies assessed the impacts of several patient- and tooth-related independent variables on tooth loss. The most frequently assessed patient- and tooth-related variables were smoking, age, gender and PD respectively.

### Reporting of logistic regression analysis

Eighty-three (23%) entries indicated that criteria had been met. The “coding of independent variables” was reported most frequently \( n = 20 \) (83%) articles], followed by the “statistical programme” \( n = 19 \) (79%)), “prevention of overfitting” \( n = 15 \) (63%) and “main effect estimates, 95% confidence interval, and statistical significance” \( n = 14 \) (58%). The “presence of multiple regression equation or table” was reported adequately in seven (29%) articles, the “fitting procedure” item was reported adequately in six (25%) articles, “testing for interaction of independent variables” were reported in two (8%) articles, and “goodness-of-fit (GOF) procedures” were reported in one (4%) article. The other criteria were not met in any article included in the analysis (Table S1). Four (17%) articles (Miyamoto et al. 2006, Matulienė et al. 2008, Bäumer et al. 2011, Ng et al. 2011) reported the inclusion of differences in patient follow-up lengths into the logistic regression model.

Regarding the trend of reporting of logistic regression models (by percentage of included articles), since 1998, less than 30% of criteria were met in all three time periods [1998–2006 \( n = 5 \) articles], 2007–2010 \( n = 11 \), and 2011–2014 \( n = 8 \)], with no difference among periods observed.

### Discussion

The findings of this systematic review demonstrate that the reporting of logistic regression analysis methods in studies assessing tooth survival in patients treated for periodontitis is not optimal. The studies included in the analysis met less than 30% of reporting criteria assessed, and the quality of reporting showed no substantial improvement since 1998. These results may be explained by the current lack of detailed guidelines for the adequate reporting of statistical assessments in studies in periodontology.

Many criteria assessed were met in few or no studies. For example, only one (4%) study reported on GOF procedures. These results are in agreement with low rates of reporting observed in other medical fields (Mikołajczyk et al. 2008, Kalil et al. 2010). GOF measures how well a model describes the response variable (Bewick et al. 2005), with a low value indicating that the covariates included in the analysis do not predict the outcome well. Assumptions based on a poorly fitting model may thus be inaccurate (Hosmer et al. 1991).

Another important issue of the logistic regression statistics is the test for multicollinearity in predictive models. Collinearity and multicollinearity occur when covariates are strongly correlated, and may distort the interpretation of a logistic regression model (Tu et al. 2005). No study included in this analysis contained a report of the test used for collinearity, although the authors of two studies reported that they assessed the collinearity of variables included in final logistic regression models. These results are in accord with findings reported in studies from other medical fields that multicollinearity was assessed in less than 0.5% of articles (Mikołajczyk et al. 2008). Most commercial software packages implement collinearity tests, such as the calculation of variance inflation factors, conditional indices, and eigenvalues. Those tests evaluate the degree of correlations among explanatory variables in the model (Tu et al. 2004).

Approximately 60% of articles included fulfilled the criterion related to prevention of overfitting, which is the excess of independent variables in the final model (Peduzzi et al.

© 2014 John Wiley & Sons A/S. Published by John Wiley & Sons Ltd
<table>
<thead>
<tr>
<th>Studies</th>
<th>Country</th>
<th>Study methodology</th>
<th>Independent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Büttner et al. (2011)</td>
<td>Germany</td>
<td>Retrospective case-series. Eighty-four patients aged 20–36 years (mean age 30.8 ± 4.1 years) and with AgP were re-evaluated after a mean period of 10.5 years of supportive periodontal therapy (SPT)</td>
<td>Sex, age, IL-1 composite genotype, diagnosis (localized or generalized AgP), SPT, smoking (never smoker versus active smoker and former smoker versus active smoker), BMI, insurance status, mean GBI during SPT, educational status (low versus high and moderate versus high), relative amount of inter-proximal bone loss (5 categories of bone loss), tooth type (molar versus pre-molar), jaw (mandible versus maxilla), FI (multi-rooted tooth with FI), abutment tooth (yes), GTR/EMD during APT or SPT, and follow-up time</td>
</tr>
<tr>
<td>Carnevale et al. (2007)</td>
<td>Italy</td>
<td>Retrospective case-series. 304 patients aged between 25 and 85 years (mean age 52 ± 11 years) were evaluated after SPT programme of 3–17 years (mean time 7.8 years)</td>
<td>Diagnosis, smoking habits, gender, age, FMPS, FMBS, time in SPC</td>
</tr>
<tr>
<td>Chambrone &amp; Chambrone (2006)</td>
<td>Brazil</td>
<td>Retrospective case-series. 120 patients with an average age of 38.9 years (20–72 years) were evaluated after 10 or more years of SPT</td>
<td>Age, gender, period of SPT, number of annual maintenance care appointments, and smoking habits</td>
</tr>
<tr>
<td>Checchi et al. (2002)</td>
<td>Italy</td>
<td>Retrospective case-series. 92 patients with a mean age of 45 years and a range of 37 years (28–65 years) treated for chronic adult periodontitis and observed over a mean period of 6.7 years under SPT</td>
<td>Categories of compliance in the SPT (categories A, B, and C). Not reported whether more variables were included in the initial model.</td>
</tr>
<tr>
<td>Costa et al. (2012)</td>
<td>Brazil</td>
<td>Retrospective cohort. 75 regular complier (RC) and 89 erratic complier patients with age range from 18 to 62 years were examined after 3 years of SPT</td>
<td>Percentage of BOP (unit increase of 1% of sites with BOP), number of periodontal pockets with PD&gt;5 mm (increase in one pocket), number of missing teeth (increase in one missing tooth), bone loss in relation to age (increase of 1% of bone loss), smoking (yes/no), and diabetes (yes/no)</td>
</tr>
<tr>
<td>Costa et al. (2013)</td>
<td>Brazil</td>
<td>Retrospective case-series (derived from Costa et al. 2012). 92 diabetic patients with age range from 22 to 71 years were examined after 5 years of SPT</td>
<td>Sex (male/female), age (4 levels), education level, cohabitation status (companion/no companion), number of PMT visits, BMI, HbA1c (% ≥ 6.5, HbA1c (%) &lt; 6.5, individuals without diabetes (reference group), smoking status (smoker, former smoker, no smoker), PD 4 to 6 mm in ≥ 30% of sites, smoking status by HbA1c ≥ 6.5% interaction, duration of DM (&gt;10 years), BOP (in &gt;30% of sites), PDE ≥30% of sites, and CAL ≥ 3 mm in 30% of sites</td>
</tr>
<tr>
<td>Costa et al. (2014)</td>
<td>Brazil</td>
<td>Retrospective cohort. 96 regular complier (RC) and 116 erratic complier patients with age range from 23 to 70 years were examined after 5 years of SPT</td>
<td>Gender (male/female), age (4 levels), cohabitation status (companion/no companion), BMI, BOP (in more than 30% of sites), CAL ≥ 3 mm in 30% of sites, irregular compliance, male gender, smoking (smoker, former smoker, and non-smoker), diabetes, PD between 4 and 6 mm up to 10% of sites</td>
</tr>
<tr>
<td>Ekuni et al. (2009)</td>
<td>Japan</td>
<td>Retrospective case-series with 25 patients who had received supportive periodontal treatment for 5–16 years following non-surgical periodontal treatment at a university hospital</td>
<td>Initial vertical bone loss, crown-root ratio, initial deepest CAL, initial BOP, initial amount of bone loss (%), history of repeated periodontal abscess formation during SPT (Y/N), initial deepest PD, initial tooth mobility, multi-rooted tooth</td>
</tr>
<tr>
<td>Faggion et al. (2007)</td>
<td>Germany</td>
<td>Retrospective case-series. 198 patients with mean age of 47.58 ± 10.42 years complying with SPT over an average of 11.8 ± 2.3 years</td>
<td>DM, CHD, infectious diseases, allergies, coagulation disorders, PD, tooth type, approximal plaque index, SBI, SPT frequency, alveolar bone level, tooth mobility, root type (multi versus single-rooted), and tooth vitality at baseline examination.</td>
</tr>
<tr>
<td>Fardal et al. (2004)</td>
<td>Norway</td>
<td>Retrospective case-series. 100 patients with an average age of 46 years, range: 25–69 years, who had comprehensive periodontal treatment and attended for 9.8 (SD: 0.7), range: 9–11 years of maintenance care</td>
<td>Gender, age, smoking, maintenance, overall health status, and family history</td>
</tr>
<tr>
<td>Fardal &amp; Linden (2008)</td>
<td>Norway</td>
<td>Retrospective case-control study. 54 patients (27 with refractory periodontitis) with an average age of 48.5 (SD 10.0, range 21–71) years were followed up for 13.4 (SD 3.3, range 8–19) years</td>
<td>Not clearly reported</td>
</tr>
<tr>
<td>Studies</td>
<td>Country</td>
<td>Study methodology</td>
<td>Independent variables</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------</td>
<td>--------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Leininger et al. (2010)</td>
<td>France</td>
<td>Retrospective case-series, 30 patients with mean age of 51.0 years (22–67), treated for periodontitis and under SPT for 6–12 years</td>
<td>Periodontal risk assessment diagram surface:&lt;br&gt;Age, marital status, diabetes, smoking (4 levels), alcohol use, family income (2 levels), BOP in more than 30% of sites, PD≥4 mm in more than 30% of sites, gender and PD between 4 and 6 mm up to 10% of sites&lt;br&gt;Relative bone height, teeth with root abnormality, ≥25% of sites with bleeding index &gt; 2; sessions of supportive periodontal therapy &gt; 1 per year&lt;br&gt;Levels of disease severity.</td>
</tr>
<tr>
<td>Lorentz et al. (2010)</td>
<td>Brazil</td>
<td>Retrospective case-series. 50 compliers individuals with mean age of 44.9 ± 9.5 years (range 18–74 years) who were included in a programme of SPT, and monitored during 12 months</td>
<td>Disease score:&lt;br&gt;Maximum PD: PD≤3 mm, PD=4 mm, PD=5 mm, PD=6 mm, PD≥7 mm, Maximum furcation class I/II/III, Mobility grade I/II/III, Maximum CAL: CAL≤4 mm, CAL=5 mm, CAL=6 mm, CAL≥7 mm, CAL≥10 mm, FMBS≥30% versus &lt;30%, diagnosis: level 2 versus level 1, SPT: ≥10–15 years versus &lt;10 years, SPT: ≥16 years versus &lt;10 years, smoking (2 levels), DM, university versus private practice, gender (male/female)&lt;br&gt;Age (45-54, 55-64), age (35-44, ≥65), gender (male/female), diabetes, smoking, PD≥5 mm (per one pocket more), risk at T1 (2 levels), SPT: ≥10 years versus &lt;10 years, Compliance: erratic versus full, Age at T1 (per 1 year older), BOP (per 1% increase in bleeding sites per patient), Bone loss (age per 1% more bone loss in relation to patient’s age)&lt;br&gt;Level of compliance (complete or erratic)</td>
</tr>
<tr>
<td>Lv et al. (2013)</td>
<td>China</td>
<td>Retrospective case-series with 85 AgP patients who were included in a programme of SPT for 3–11 years</td>
<td>Level of compliance (complete or erratic)</td>
</tr>
<tr>
<td>Martin et al. (2010)</td>
<td>USA</td>
<td>Retrospective case-series. 776 patients with mean age of 46.0 ± 10.5 years treated and maintained by nine periodontists in their clinical practices (100 patients in each practice) (mean period of 13.2–7.0 years (range: 3.0 to 32.5 years))</td>
<td>Level of compliance (complete or erratic)</td>
</tr>
<tr>
<td>Martin et al. (2011)</td>
<td>USA</td>
<td>Retrospective case-control study with a historic comparison population (N = 443) and periodontally treated population (N = 653) followed up for more than 10 years</td>
<td>Level of compliance (complete or erratic)</td>
</tr>
<tr>
<td>Matuliene et al. (2008)</td>
<td>Switzerland</td>
<td>Retrospective case-series. 172 patients between 14 and 69 years of age (mean 45 ± 11 years) under SPT for 3–27 years (mean 11.3 years)</td>
<td>Level of compliance (complete or erratic)</td>
</tr>
<tr>
<td>Matuliene et al. (2010)</td>
<td>Switzerland</td>
<td>Retrospective case-series with 172 patients under SPT maintenance for 9.5 ± 4.5 years (same sample from Matuliene et al. 2008)</td>
<td>Level of compliance (complete or erratic)</td>
</tr>
<tr>
<td>Miyamoto et al. (2006)</td>
<td>Japan</td>
<td>Retrospective case-series with 505 compliant and non-compliant patients maintained under SPT for at least 10 years</td>
<td>Level of compliance (complete or erratic)</td>
</tr>
<tr>
<td>Ng et al. (2011)</td>
<td>Singapore</td>
<td>“Longitudinal retrospective study and a clinical recall cross-sectional study”. The longitudinal study is a retrospective case-series with 273 compliant and non-compliant patients (mean age 44.7 years, range 19–80) who were maintained under SPT for at least 7 years</td>
<td>Level of compliance (complete or erratic)</td>
</tr>
<tr>
<td>Pretzl et al. (2008)</td>
<td>Germany</td>
<td>Retrospective case-series with 100 compliant and non-compliant patients (mean age 46.5 ± 10.3) were assessed after at least 10 years of SPT maintenance (10 years ± 6 months)</td>
<td>Level of compliance (complete or erratic)</td>
</tr>
<tr>
<td>Ravald &amp; Johansson (2012)</td>
<td>Sweden</td>
<td>Retrospective case-series with 64 patients (mean age: 64 years, SD: 8.3) reassessed after 11–14 years (mean 12.5 years) (SPT of at least 2 years)</td>
<td>Level of compliance (complete or erratic)</td>
</tr>
<tr>
<td>Tonetti et al. (1998)</td>
<td>Switzerland</td>
<td>Retrospective case-series with 273 patients (mean age 52 years ±14) with an observation period of 67 ± 46 months (range 5 months to 23 years)</td>
<td>Level of compliance (complete or erratic)</td>
</tr>
</tbody>
</table>
**Table 3.** Independent variables

<table>
<thead>
<tr>
<th>Independent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Smoking, initial tooth prognosis, tooth type, participant's compliance with suggested maintenance visits, and Acceptability of the quality of restorations, age, sex</td>
</tr>
</tbody>
</table>

Only a small part of articles reported whether the follow-up of patients was taken into consideration when developing the logistic model. When performing and interpreting the model, special attention should be paid to censoring, i.e. when the outcome or observation is partially known (Bewick et al. 2005). Censoring may occur due to incomplete follow-up (e.g. when a patient dies or withdraws before the end of the study) or a patient does not experience the event of interest, i.e. tooth loss. Survival analyses techniques are focused on time to event data and, therefore, they may be more indicated to evaluate data when censoring is present (Bewick et al. 2005), although some other models such as conditional logistic regression might be applied to such cases (Rabinowitz et al. 2000).

No article reported any information about the external validation of the logistic regression models. To understand the clinical validity of a prediction model, its performance should be evaluated with data from patients not included in the dataset used to develop the model (Collins et al. 2014). The model should be evaluated against its accuracy to predict the outcome through calibration, i.e. the agreement between the prediction from the model and the observed outcome (Collins et al. 2014). The other main component to evaluate the performance of the model is the discrimination, i.e. the ability of the model to separate the patients with and without the evaluated outcome (Collins et al. 2014). Discrimination may be quantified using approaches which estimate the area under the Receiver Operating Characteristics (ROC) curve (AUC), known as a “C-statistic” (Steyerberg et al. 2010, Uno et al. 2011). Higher values of AUC indicate that the logistic model discriminates well (Bewick et al. 2005).

**Limitations of the study**

The assessment of a logistic regression model is not always straightforward. For example, the testing of a linearity assumption depends on the type of non-linearity against which the assumption is measured; in a model including age and age squared, the finding that the age-squared variable does not have a significant effect may lead to the assumption of a linear relationship. However, this assumption may not be true, as the correct functional form of the variable may be log age and the model may have too little or too much statistical power to reject the linearity assumption. Thus, some criteria assessed in this review are not expected to be applicable to all studies using logistic regression analysis.

**What this study adds to the available evidence**

To our knowledge, this study is the first to assess and discuss the reporting of statistical approaches, specifically logistic regression models, in studies published in dentistry. This lack of information on statistical assessments hinders the reader’s ability to comprehensively understand the validity of reported findings. In the case of the assessment of risk factors for tooth loss in patients who have received periodontal treatment, the lack of detailed reporting about logistic models may lead to the reporting of insufficiently accurate information about whether a specific risk factor is correlated with a higher probability of tooth loss. Thus, our confidence that eliminating some risk factor identified in the model improves tooth survival may be reduced (e.g. in the case of the use of surgical means to reduce PD, a potential risk factor for tooth loss identified in the model). Furthermore, the findings of this systematic review may help clinicians to identify topics related to logistic regression reporting to which they should pay particular attention when reading published studies.

**Recommendations for researchers**

Several measures should be taken to improve the quality of reporting of logistic regression models in periodontology. Editors of periodontology journals should require that authors meet the following minimum standards for the publication of
studies involving logistic regression analysis. These standards should include the report of following procedures as suggested by Kalil et al. (2010):

- Odds ratios and 95% confidence intervals
- Overfitting prevention procedures
- GOF procedures
- Multicollinearity testing
- Interaction testing
- Internal validation (if sample size allows)
- Independent variable selection
- Statistical programme used.

Detailed reporting of a statistical model may considerably increase the length of an article. The majority of scientific journals currently provide space in an online format (e.g. appendices) to accommodate relevant information that cannot be reported in the printed version of the paper. Editors of periodontal journals should also adopt and endorse some general guidelines for reporting statistics in articles of clinical trials such as the Statistical Analyses and Methods in the Published Literature (SAMPL) guidelines (Lang & Altman 2013). These guidelines may improve the overall quality of reporting of common statistics in periodontal articles and also prevent reporting errors (Lang & Altman 2013).

Future research

Standards of reporting of other statistical methods used to assess risk factors for periodontitis development and tooth loss, such as linear regression (Palmer et al. 1999) and survival analysis (McGuire & Nunn 1996), should be evaluated using an approach similar to that of this study. The present project was focused on evaluating the standards of reporting logistic regression models. Future projects should address whether the regression models were correctly applied in periodontal research.

Conclusions

The reporting of logistic regression analysis methods in studies of tooth survival in patients treated for periodontitis is not optimal. High-quality reporting of statistical methods is critical for readers’ comprehensive understanding of published studies. The results of this systematic review can aid researchers’ reporting of logistic regression analyses in periodontology.

References


**Supporting Information**

Additional Supporting Information may be found in the online version of this article:

**Appendix S1.** List of excluded papers with reasons for exclusion.

**Table S1.** Methodological quality of logistic regression analysis in studies assessing tooth survival in periodontitis treated patients (1: criterion met; 0: criterion not met; 2: unclear).

**Address:**

Clovis Mariano Faggion Jr
Faculty of Dentistry
University of Münster, Waldeyerstraße 3048149 Münster
Germany
E-mail: clovisfaggion@yahoo.com

---

**Clinical Relevance**

**Scientific rationale for the study:** Comprehensive description of logistic regression analysis is critical for an adequate understanding of the validity of results. No previous study has assessed the reporting of logistic regression analyses in periodontology. **Principal findings:** Detailed reporting of information about logistic regression models is lacking in studies assessing risk factors for tooth loss in patients who have received periodontal treatment. **Practical implications:** The present findings may guide authors of clinical periodontal research in the adequate reporting of logistic regression methodology, and enhance readers’ understanding of the validity of reported results.